# INDIAN STATISTICAL SYSTEM

COMMEMORATIVE VOLUME 1

GOLDEN JUBILEE
OF INDIAN INDEPENDENCE

DEPARTMENT OF STATISTICS
MINISTRY OF PLANNING AND
PROGRAMME IMPLEMENTATION
GOVERNMENT OF INDIA
NEW DELHI

# INDIAN STATISTICAL SYSTEM

## Golden Jubilee of

## Indian Independence

सत्यमेव जयते

Department of Statistics

Ministry of Planning and Programme Implementation

Government of India

June 1998

Commemorative Volume

1

*Professor P. C. Mahalanobis, father of Indian statistics*
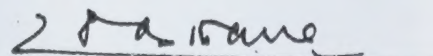*(29 June 1893 - 28 June 1972)*

# FOREWORD

The Government of India is currently celebrating the Golden Jubilee Year of the Nation's Independence. A series of events to commemorate the 50th Anniversary of our Independence has been drawn up by the Ministries and Departments of the Government. Amongst the various programmes included in the calender of events prepared by the Department of Statistics is the issue of a special publication in the form of two "Golden Jubilee Volumes".

I am happy to note that in response to our request 32 articles have been received for this publication, a number of them from Indian statisticians currently working abroad in foreign universities and institutions. The authors have given useful suggestions on the direction in which the Department of Statistics must work and their views on the future of statistics in this country. I hope that this publication will be useful to a wide circle of readers including those who are not directly connected with the official statistical system. The suggestions /recommendations emanating from the views expressed by different authors, especially the non-departmental outsiders will go a long way in formulating a concrete plan of action and providing a platform for the Department of Statistics for launching into the twenty-first century.

The editorial board constituted with Prof. Atindra Mohan Goon, Prof. Nikhilesh Bhattacharya and Prof. Arijit Chaudhuri, all eminent statisticians by themselves, have made significant contributions in editing the articles received. In addition to the members of the editorial board, I would like to congratulate all the other members of this Department who have helped in the successful publication of this volume.

(M.D.Asthana)
Secretary
Department of Statistics
Government of India

*Professor Mahalanobis with Prime Minister Nehru*

# PREFACE

The late S. Subramanian presented at the International Statistical Conference held in Calcutta in 1951, a brief history of the organization of Indian official statistics up to the attainment of the country's independence from British rule in 1947. He refers to the *Arthasastra* and the *Ain-i-Akbari* to remind us that collection of numerical data for purposes of statecraft was not unknown in medieval or even ancient India. The British period was marked by a consolidation of efforts for the collection of socio-economic data. Just a year after the Battle of Plassey, in 1758, Frankland brought out his *Statistical Report on the Pergunnahs Southward of Calcutta* (presumably meaning the area now covered by the two 24-Parganas districts of West Bengal). Subramanian also tells us of the recommendations of various committees and commissions that went into the question of data-collection and the changes made in the system following these recommendations till 1947.

The present commemorative volume is concerned with developments in the subsequent decades, which resulted mainly from the association of Professor P.C. Mahalanobis with the official statistical system. We can surely look back at this period with a sense of pride, for the changes that have taken place in the system during this period are due entirely to an Indian initiative and are in conformity with modern statistical theory and practice.

Indeed, the Indian system has been a role model for those of many other countries of the world. Yet, it is advisable to look critically at the achievements of the system, detect shortcomings and lacunae and think out means for improving the system.

The present commemorative volume was conceived in this spirit of reappraisal of the system while celebrating the Golden Jubilee of the nation's independence. Leading Indian academics, working within the country or residing abroad, as well as statisticians who were at the helm of the system in the past or occupy responsible positions at present, were requested to contribute to the volume giving their valued opinions and their 'vision of the future of statistics in the country'. The contributions of all those who have responded are included in this volume. It also includes an excellent article received from Dr. C. M. Jarque, President, National Institute of Statistics, Geography and Informatics, Mexico, which is based on a lecture delivered by him at the 1991 conference of the International Statistical Institute and dedicated to the memory of Professor Mahalanobis.

The articles have been placed under four distinct categories : (a) those of a general nature which give the authors' assessment of, and suggestions for improvement of, the work of the various wings of the Department of Statistics; (b) those that address themselves to some specific activities (like the Annual Survey of Industries or collection of Industrial Statistics) and suggest ways of improvement of those functions; (c) articles that are concerned with methodological issues, e.g. how the NSS is now carried out and what are the drawbacks of the current system, how they can be corrected, etc.; and (d) miscellaneous articles (e.g. articles suggesting how the present system of training of statisticians can be geared to the needs of the official statistical system). Within each group, the articles are arranged in alphabetical order of the authors' surnames.

As editors of the volume, we have mostly left the writings unchanged, attending simply to such minor matters as uniformity of spelling, uniformity of layout and arrangement of paragraphs in a proper sequence. Wherever substantive changes were found necessary, the authors were contacted and their prior consent received to changes suggested by us. In one or two places short editorial

notes have been appended to convey the stand of the DoS on some issues that might appear somewhat controversial and to bring to light some of the recent developments.

Some additional material has been prepared by the DoS officials and incorporated in the volume. There is right at the beginning an overview of the functions of the three wings of the DoS, viz. the NSSO, the CSO and the DoS Computer Centre. This should help the reader form a good idea of the evolution of the official statistical system and its wide range of activities. The contributed articles are followed by a summary highlighting the major recommendations of the authors; this should be helpful to the DoS authorities in further improving the mode of functioning of the DoS. Lists of the publications - reports and notes - brought out by the NSSO and CSO in recent years are supplied to help the readers, especially researchers, who may feel inclined to draw upon these two important sources of data. Lists of important research papers based on NSSO, CSO or RGI data that come out in the two prestigious journals *Economic and Political Weekly* and *Indian Labour Journal* have also been prepared by the DoS officials and included in the volume. These should give a good idea about the extent to which such official data are being actually utilized by our researchers.

We sincerely hope that this volume will be read with interest and will enable the DoS to find a proper direction for its functioning in the years to come.

It has been a privilege and a delightful experience for us to have been associated with the publication of this volume. While acknowledging our satisfaction, we should also put on record our appreciation of the excellent cooperation received from Dr. V. Saha, Director, DPD & SDRD, NSSO and his colleagues, that has enabled us to prepare the press copy of the volume almost in a month's time.

**- The Editors**

# Contents

**Methodological Issues :**

**Specific Activities of the DoS :**

**Miscellaneous Topics :**

# FUNCTIONS OF THE DEPARTMENT OF STATISTICS
## An Overview

In India, collection of statistical data which were of use to the sovereign was practised even in ancient times. More recently, the Mughals had a system of collection and compilation of crop statistics to help them in land revenue collection. With the progress of British rule these systems died out; the British, in course of time, created their own data-generating systems to serve their specific ends. These systems of data collection did not develop into an integrated or well-coordinated statistical system during colonial rule and their coverage was largely limited to a few specific fields like trade and commerce, production of selected industries, population, some basic crop statistics and livestock.

Independence in 1947 ushered in an era of economic planning and emphasised the necessity of a strong database covering a variety of social and economic topics. The system of data collection left behind by the colonial rulers was found to be far from adequate to meet this immediate demand. It did not even provide the basic data required for estimation of national income which was so vital for assessing performance and progress of the economy. Data reflecting the condition of the vast majority of the population languishing in a state of chronic hunger were hard to come by. There was no provision for generating employment statistics for the large section of population engaged in agriculture and other unorganised sectors of the economy. The immediate task, therefore, was to set up a statistical system capable of filling the large gaps in the data essential for formulating effective economic plans.

At the instance of Prime Minister Jawaharlal Nehru, a modest beginning was made in 1949. The National Income Committee was appointed by the Government of India to work out a reliable method of estimating national income. On its recommendation, the Directorate of National Sample Survey was set up in 1950 to collect essential statistics relating to socio-economic characteristics and agricultural production. In the following year, the Central Statistical Organisation was formed. In short, the foundations of a modern statistical system were laid in those early years of Independence. Since then, the sustained efforts of academics and official statisticians have seen the collection of statistics in India gradually mature into one of the most comprehensive statistical systems of the developing world.

Today, the Department of Statistics(DoS) of the Government of India is the apex body in the official statistical system of the country. It comprises the Central Statistical Organisation (CSO), the National Sample Survey Organisation (NSSO) and the Computer Centre (CC) of the Department. The Department is also responsible for all policy matters of the Indian Statistical Institute (ISI), an autonomous body of international repute, whose contributions in the development of statistical theory and practice are acclaimed worldwide.

## Creation of the National Sample Survey Organisation

The National Income Committee noticed large gaps in the statistical database of the country and recommended that improvements be brought about in the sphere of information as well as in the sphere of material production. Professor P.C. Mahalanobis, who, as head of the Indian Statistical Institute, had been directing the large-scale sample surveys of crop acreage and yield for several years, proposed to the Government of India the setting up of a permanent survey organisation to conduct national sample surveys to assist in socio-economic planning and policy-making. Accordingly, the National Sample Survey (NSS) came into being in 1950 and the first round of the NSS, covering rural India, was conducted during October 1950 - March 1951.

## Formation of the Central Statistical Organisation

For the purpose of coordination of statistical activities of the different Ministries of the Government of India and the State governments and for promotion of statistical standards, the Central Statistical Organisation (CSO) was created in May 1951. Since then the CSO has been coordinating the statistical activities in the country, including laying down and maintenance of statistical norms and standards and providing liaison with Central, State and international statistical agencies. The CSO also shoulders the responsibility of preparation of national accounts, compilation and publication of industrial statistics, conduct of economic census and surveys, middle class family income and expenditure surveys, training, compilation of price statistics, human development statistics, environment statistics and dissemination of statistics on various socio-economic aspects of national life.

## Setting up of the Computer Centre

The Computer Centre (CC) was set up in 1967 as an attached office of the Department of Statistics to cater to the data processing needs of the Department and other Departments of the Union government. From its very inception, the Centre has effectively performed the pioneering task of building up a database for many organisations of the Government of India and played a vital *rôle* in imparting intensive training in systems analysis, data processing and consultancies. The Centre's major responsibility relates to processing of data collected by the NSSO through various socio-economic surveys, follow-up surveys of the Economic Census, and the Annual Survey of Industries conducted by the NSSO and the CSO.

## Constitution of the Indian Statistical Service

The wealth of information that the statistical system of the country was required to generate needed services of a cadre consisting of a large group of trained professionals for its proper handling and analysis. With this end in view, the Indian Statistical Service (ISS) was constituted in November 1961 by pooling together posts in statistical discipline existing in various ministries which were offered for inclusion in the ISS.

## Contribution of the NSSO to the National Database

The NSS came into existence in 1950 and immediately proceeded to conduct multi-subject household enquiries, designed to fill gaps in the data required for planning, in randomly selected villages and urban blocks spread over the entire geographical area of the country. With increasing demand for reliable estimates at State level and regional level, the need for larger samples has been increasingly felt, and efforts made to enlarge the sample size over the years. At present the NSS covers, in each round, a sample of about 12,000 to 14,000 villages and blocks at the all-India level in the "Central sample" (covered by the NSSO) and an independent sample of about 14,000 to 16,000 villages and blocks at the all-India level through "State samples" (covered by various States/ Union Territories). The field work for the Central sample is undertaken by a team of highly trained investigators of the Field Operations Division of the NSSO.

NSS data on socio-economic surveys are regularly released through the quarterly publication *Sarvekshana* issued by the Department of Statistics.

**Household Consumption** The all-India household consumer expenditure survey started with the first round ( 1950-51) of the NSS, giving the first all-India picture (though then restricted to the rural sector) of levels and patterns of household consumption. The NSS enquiries on consumer expenditure are now the main sources of data on level of living of the Indian population.

**Employment and Unemployment** The survey on employment and unemployment started with the 9th round (1955-56) of the NSS to assess the volume and structure of employment and unemployment. Thereafter, the NSS has tried out different concepts and definitions for its surveys on employment and unemployment in different rounds. The Dantwala Committee Report (1970)

gave a firm conceptual framework for such surveys. The first survey based on the concepts and definitions recommended by this committee was conducted in the 27th round (1972-73) of NSS. Since then the NSS survey on employment and unemployment has become a quinquennial feature. Four such surveys have been conducted - in the 32nd (1977-78), 38th (1983), 43rd (1987-88) and 50th (1993-94) rounds of the NSS - with the same basic approach to measurement of employment and unemployment.

**Manufacturing and Trade**   In 1951, the Directorate of NSS was made responsible for the collection of data on the factory sector through the Sample Survey of Manufacturing Industries (SSMI) started in order to extend the coverage of the Census of Manufacturing Industries (CMI). Now statistics on the entire manufacturing activity of the country - both organised and unorganised - are collected by the NSSO.

(a) *Organised Manufacturing*: While the Field Operations Division of the NSSO collects data through the Annual Survey of Industries (ASI), the CSO is responsible for designing the surveys and publication of the results.

The construction of Input-Output Transaction Tables for the Indian economy is largely based on ASI data. The tables provide the basic ratios required for application of the Input-Output Technique, which has a wide range of uses from the analysis of economic structure and formulation of economic programmes, to the setting of targets for production, investment, employment, etc., at the macro level.

(b) *Unorganised Manufacturing and Trade*: Nearly one-third of the national income emanating from the entire manufacturing activity of the country comes from the unorganised segment. The NSSO conducted, between 1950-51 and 1974-75, surveys on the unorganised segment of manufacturing trade and transport. Based on these, a new programme of data collection from the unorganised sector has been launched. Under this scheme, an Economic Census (EC) is conducted periodically and is followed up by sample surveys. The follow-up surveys have been specially designed to fill the data gaps relating to the unorganised segment of the non-agricultural sector of the economy.

**Agriculture** *Agrarian Structure*:  In 1954-55, the NSS conducted the first household survey of land holdings in its 8th round. This historic enquiry threw up data on the size distribution of land holdings at a time when land reform legislation was being enacted in different States. The NSS enquiries on land and livestock holdings, the last having been conducted in 1992, provide the data required to assess the changes in the distribution of land over time.

*Crop Area and Yield*: Crop estimation surveys covering principal crops, started by the NSS in 1957-58, helped to put crop statistics in India on a sound statistical footing. The crop surveys of the NSS were able to build on its exploratory work in earlier rounds which helped it to acquire in-depth knowledge of the inter-regional diversity of cropping patterns prevalent in India.

Currently, the collection of data on area under different crops is done by the State machinery on the basis of complete field-to-field inspection. To estimate the yield rate, crop-cutting experiments are conducted by the State agencies. In both area and yield estimation, the NSSO Field Operations Division plays a supervisory *rôle*.

**Human Resource Development**  Data on literacy, education, health and morbidity are essential for evaluation of the need for government action in each of these areas. The NSSO provides periodic estimates of (i) literacy, (ii) school enrolment, (iii) utilisation of educational services, (iv) general morbidity, (v) maternity and child care, and (vi) utilisation of medical services.

*Disability*: The latest survey for which data are available was conducted in 1991. The survey aims to estimate the total number of physically disabled persons in the country, the number of persons suffering from more than one type of disability (visual/hearing/speech/ locomotor) and rates of prevalence and incidence of different types of disability.

*Utilisation of the Public Distribution System*: The data collected enable estimation of percentage of population using the public distribution system(PDS), separately for different social and economic categories and for different commodities coming under the PDS.

**Housing Condition** The NSS started collection of data on condition of dwelling units, and basic housing amenities available to them, from its 7th round (Nov.'53 - Mar.'54). However, comprehensive independent surveys on housing condition to provide estimates at the State/U.T. level for both rural and urban areas of the country were first undertaken in 1973-74, and then in 1988-89 and 1993. Information on drinking water, sanitation and hygiene are collected by the NSSO sometimes as part of the housing condition surveys and sometimes separately.

**Debt and Investment Surveys** The first such survey was conducted by the Reserve Bank of India (RBI) in 1951-52 and christened the "All-India Rural Credit Survey". This and the subsequent one, carried out in 1961-62 again by the RBI were limited to rural areas only. However, in 1971-72, this survey was conducted along with the Land Holdings Survey of the 26th round of the NSS and covered the urban areas too, a practice which was further continued in the subsequent 37th (1982) and 48th (1992) rounds.

**Prices** A continuous series of retail and wholesale prices statistics is available in India for a fairly long time. The rural retail prices data are being collected by the NSS since the early 50's - from the 5th round (Dec.'52 - Mar.'53) onwards. In addition, the NSS collects data on retail prices from a number of urban centres. The rural retail prices are used by the Labour Bureau to construct consumer price index (CPI) numbers for agricultural labourers, and the urban retail prices are used by the CSO to construct CPI numbers for non-manual workers.

**Recent Developments in the NSSO** The 1990's have been marked by modernization of the work of the NSSO on a number of fronts. A significant step taken in the recent past is the use of palm-top computers for collection of data in the field. These were introduced in the 52nd round (1995-96) on an experimental basis with the objective of reducing the existing time lag in release of suevey results significantly.

Computerisation of the sample selection process, use of modern software packages in design of schedules and instruction manuals and in-house development of generalised software packages for data entry and data validation have eliminated several sources of error and improved the work environment in the NSSO's three main divisions: SDRD, DPD and FOD. Attractive and lucid presentation of results is now achieved by the use of computers. The NSS reports are now offered for sale to the public as soon as they are released. Reports are sold both in the form of hard copy and on magnetic media. Also on sale are raw (household-level) data, containing a wealth of information that scholars in various disciplines need for research. These are available on magnetic media at a reasonable cost.

The increased use of computers has led to great advances in the work of data processing in the last few years. The entire job of data processing, from data entry to table generation, is now done at the Data Processing Division of the NSSO. Intensive study of sources and incidence of error has led to the development of appropriate software packages for identification of errors, particularly "howlers", and cleaning of data.

With these improvements, the NSSO was able to draft all the reports based on the surveys conducted up to the 51st round (1994-95) by the end of March 1998. Draft reports on surveys conducted during the 52nd round (1995-96) and the 53rd round (1997) are under preparation. Data processing for surveys up to the 53rd round (inclusive) was completed in March 1998. Data processing for the 54th round (Jan-June 1998) is in progress simultaneously with field work.

## Achievements of the CSO

**Coordination of Statistical Activities** The CSO coordinates statistical activities undertaken by the Central government departments, State governments and international agencies. To this end, the CSO participates in the interdepartmental meetings of Working Groups, Technical Advisory

Committees and Standing Committees set up by Central ministries and departments in subject areas of their concern to ensure adoption of statistical standards on a uniform basis, minimise duplication of efforts, and promote upgradation of quality and timely release of results. It also participates in the meetings of the high-level Statistical Coordination Committees of State governments which review activities relating to generation and dissemination of data within the States.

Once in two years, the CSO organises a Conference of Central and State Statistical Organisations; this provides a forum for exchange of views and experiences concerning statistical development activities in the country. The CSO also provides secretarial assistance to the National Advisory Board on Statistics (NABS), the apex body set up by the Government of India to provide an overall perspective for statistical development and to identify priority areas.

**National Advisory Board on Statistics** To provide guidance for policy issues and for ensuring effective coordination of all statistical activities of the government, the DoS constituted the National Advisory Board on Statistics (NABS) in September 1982. The NABS examines the issues relating to the field of data collection, identification of gaps in the database, duplication in collection of data, storage and retrieval of data, etc., and advises the government for improvement in the quality of statistics and their timely release. The NABS also serves as a clearing house of large-scale statistical operations.

Since its initial constitution in 1982, the NABS has been re-constituted six times. It was recently re-constituted in 1998 with the Minister of State for Planning and Programme Implementation as the chairman of the Board. The Board consists of eminent statisticians from research institutions and representatives of Central Ministries/Departments and State Directorates of Economics and Statistics.

**National Accounts** The national accounts system is intended to present a unified statistical picture of the economy, with details of inter-relations among sectors, among different types of economic activities and among regions, as are required for a complete understanding of the behaviour of the economy.

Efforts made in the pre-Independence era to compute the nation's income were sporadic, based on aggregated information, and, because of limited resources, relied heavily on unwarranted simplifying assumptions for lack of requisite data.

*First Estimates of National Income*: Recognising the need to provide regular estimates of national income, the Government of India set up, in 1949, an Expert Committee, known as the National Income Committee, with Prof. P.C.Mahalanobis as Chairman. The National Income Unit, formed to assist this Committee, was, on the Committee's recommendations, transferred to the Ministry of Finance to take charge of the work on a regular basis. The first official estimates of the national income, prepared by the CSO at constant prices with base year 1948-49, as well as at current prices, were brought out in 1956 in the publication *Estimates of National Income*. The coverage of the National Accounts Statistics was gradually extended to incorporate, successively, estimates of private consumption expenditure, capital formation and savings, factor incomes, consolidated accounts of the nation and detailed accounts of the public sector. In 1978, a White Paper was published introducing another change in base year, this time to 1970-71, for the constant prices series. Estimates of this new series were published in the *National Accounts Statistics*, 1978.

*New series of National Accounts*: In February 1988, the CSO released a new series of national accounts aggregates with 1980-81 as base year. Various methodological improvements were carried out in preparing the 1980-81 series. Efforts were made to use as much current data as possible to replace the proportions based on the old benchmark surveys. Special mention should be made of data collected under the programme of follow-up surveys of the Economic Census.

The 1997 issue of the annual publication *National Accounts Statistics* includes estimates of national product, domestic product, private final consumption expenditure, savings, capital formation, capital stock, public sector transactions and consolidated accounts of the nation for the

period 1980-81 to 1994-95. It also includes a special statement on domestic product from livestock sector for the same period.

For a more recent year, viz. 1996-97, the CSO has published "quick" estimates of national income and related aggregates. Moreover, advance estimates for 1997-98 were released in February 1998.

**Collection of Industrial Statistics** Systematic collection, compilation and publication of basic industrial statistics in this country started only after the enactment of the Industrial Statistics Act, 1942. Under this Act, a Census of Manufacturing Industries (CMI) was started in 1946.

Coverage of the CMI was limited to 29 out of 63 groups into which manufacturing had been classified and, further, only to factories which employed 20 or more workers and used power. Geographically, too, the CMI did not cover the entire country. A Sample Survey of Manufacturing Industries (SSMI), covering all 63 industry groups, was started in 1950 with reference year 1949. In 1952, the work of SSMI was passed on to the National Sample Survey and its coverage was extended to all factories employing 10 or more workers if using power and 20 or more if not using power.

The work of the CMI and SSMI were independent and overlapping. To avoid duplication, the two were integrated into the Annual Survey of Industries (ASI), introduced from 1960 with reference year 1959. Following the enactment of the Collection of Statistics Act (1953), the Chief Director, NSS, was appointed in 1960 as the Statistics Authority under this Act, empowered to call for whatever information on industrial statistics was required in the country. Since then, the ASI is being conducted regularly, its field work being entirely done by the NSS.

The data collected in the ASI relates to value added, capital, employment, emoluments, output, consumption of raw materials, fuel, lubricants, etc., and other characteristics of factories and industrial establishments. Besides all factories registered under Sections 2m(i) and 2m(ii) of the Factories Act, 1948, the ASI covers *bidi* and cigar manufacturing registered under the *Bidi* and Cigar Workers Act, 1966, undertakings engaged in generation, transmission and distribution of electricity registered with the Central Electricity Authority, cold storage, water supply and certain repairing activities.

The summary results of the ASI (in 2 volumes), for the year 1994-95, were released in November 1997. As for the detailed results, the CSO has resumed publication of detailed ASI reports from 1993-94 onwards. The detailed results for 1993-94 are available both in printed form (15 volumes) as well as on magnetic media. For 1994-95, the detailed results are available on magnetic media and 3 out of 16 volumes have been released in report form. The remaining volumes of the report are being released in printed form in a phased manner.

The sample design for the ASI is being revised so as to reduce the sample size. This revised design will be adopted from ASI, 1997-98. Also, in order to reduce the response burden on the part of the manufacturing units, a much shorter schedule of enquiry is being prepared.

**Index of Industrial Production** The index of industrial production (IIP) prepared by the CSO is the most up-to-date indicator of industrial growth in the Indian economy. The index of industrial production data are released every month and are revised twice subsequently on receipt of updated data. Till recently, up to March 1998, the index was compiled with base year 1980-81.

To focus on the structural change the industrial sector has undergone since 1980-81, the base year has been shifted to a more recent period (1993-94). The revised series of IIP for the period April 1994 to October 1997 have been released in May 1998. The CSO will release only the revised series IIP for April 1998 onwards.

**Economic Census and Surveys** The CSO has been making concerted efforts to develop a sound and reliable database on unorganised non-agricultural economic sectors. To achieve this, a scheme "Economic Census and Follow-up Surveys" was launched in 1976. Economic Census (EC) provides basic details about all economic enterprises in the country. The follow-up surveys provide detailed

information on various segments. This information is used by different categories of users and also for the preparation of national accounts.

*Economic Census*: After the first EC in 1977, the second and third ECs were conducted in 1980 and 1990 along with the house-listing operations of the 1981 and 1991 Population Censuses. The fourth Economic Census is being conducted independently by the CSO in collaboration with the State Directorates of Economics and Statistics, covering establishments and own account enterprises of the entire economy except those engaged in crop production and plantation.

*Enterprise Surveys*: The follow-up surveys are the only source of information on the unorganised sector of non-agricultural production for the preparation of the national accounts. The CSO has been conducting sample surveys on different sub-sectors of the unorganised sector. In the recent past, the following sample surveys have been carried out:

> Unregistered Directory Manufacturing Establishments (1989-90)
>
> Medical, Educational, Financial, Cultural and other Services (1991-92)
>
> Mining & Quarrying and Storage & Warehousing (1992-93)
>
> Hotels & Restaurants and Transport (1993-94)
>
> Directory Trading Establishments (1996-97).

**Human Development Statistics** The CSO plays the *rôle* of a coordinator covering the fields of population, human development, manpower, employment, health, education, social welfare, etc., and maintenance of specific standards in these fields.

The CSO also compiles Human Development Statistics and publishes them in the periodical publication *Annual Statistical Abstract*. In recent years, the CSO has brought out more comprehensive publications in the field of human development, entitled *Women and Men in India, 1995* and *Selected Socio-economic Statistics in India* (1994). These provide time series data on different indicators of human development.

**Directory of Sample Surveys in India** The CSO annually brings out a publication *Sample Surveys of Current Interest in India* making available in a single document consolidated information on objectives, coverage, sampling design, procedures adopted for data collection, arrangements for data processing, cost of survey, main findings, etc., in respect of surveys reported by various governmental and non-governmental agencies in the country. The title of the publication has now been changed to *Directory of Sample Surveys in India*, the new title coming into effect from the 1995 issue onwards.

**Environment Statistics** The CSO has assumed the responsibility for development of statistical standards in environment statistics and strengthening of database in this area. A small cell has been created in the CSO for compilation of data on environment statistics under the six internationally recognised heads: Flora, Fauna, Atmosphere, Water, Land/Soil and Human Settlements. Recently, the CSO brought out (1997) a *Compendium of Environment Statistics* which gives a comprehensive picture of environmental degradation, its causes and the reasons for concern.

**Price Statistics** The Technical Advisory Committee (TAC) on Cost of Living Index was set up by the Ministry of Labour & Employment in 1954. The TAC recommended that family budget enquiries be conducted to cover the urban middle class population all over the country. Accordingly, the CSO conducted in 1958-59 the first nationwide family living survey on urban non-manual employees. The survey was conducted in 45 selected urban centres.

This survey provided the basic data required for working out the weighting diagram for the construction of cost of living index for urban non-manual employees - CPI (UNME) - which CSO has been regularly compiling on a monthly basis since 1960.

Later, to change the base year from 1960 to a more recent year, the second family living survey was conducted during 1982-83, covering a larger number of urban centres (59). The revised

CPI series with 1984-85 as base is now compiled and released every month by the CSO. Data on urban retail prices are collected from the 59 centres for about 250 items, including 15 selected service items. Information is also collected on house rent and offtake of commodities from the public distribution system.

The CPI(UNME) index has a wide range of uses: as an input to the wage and salary adjustment process; for indexation of pension and superannuation payments; government taxes, charges and contracts; and as a general measure of inflation.

Recently, the CSO has been able to substantially reduce the time lag in release of CPI(UNME). At present the time lag has been brought down to less than a month. The index for the month of April 1998, for example, was released on 27th May, 1998.

**Training**   The Training Division organises regular training courses for the officers of the Indian Statistical Service (ISS) and statistical personnel of the Central and State governments and public sector undertakings. Besides this, the CSO organises specialised training courses for officers of the ISS and those sponsored by certain international organisations like the International Statistical Education Centre (ISEC).

The subjects covered in these courses range from Official Statistics, National Accounts Statistics, and Financial & Banking Statistics to Modern Sampling Techniques, Electronic Data Processing and Software Packages. Separate courses are organised for senior, middle and junior level officers, technical staff, students and direct recruits to Grade IV of the ISS.

## Achievements of the Computer Centre

The Computer Centre, established in 1967, started with just three second-generation computers. Since then, it has developed immensely in its data processing capabilities for different types of surveys and censuses conducted by the Department of Statistics. Today, the centre possesses a main-frame computer system and a group of highly skilled and experienced EDP personnel.

**Data Processing**   Over the years, the Computer Centre has undertaken computing jobs for processing of voluminous data varying widely in structure and content. It has shouldered the responsibility of tabulation of NSS data (and been partly involved in the data cleaning effort) right from the 27th round (1972-73) onwards to the recent 50th round (1993-94). Since 1974-75, the Computer Centre has been processing the ASI data. The entire software for processing and tabulation of data of the Economic Census (EC) of 1970, 1980 and 1990 was developed at the centre. The software developed for the EC,1990 at the centre was provided to the States for processing and tabulating State-level data. The data of different enterprise surveys of the CSO are processed at the centre. The processing of monthly price data - from data entry to production of final price index for urban non-manual employees - is also done regularly at the Computer Centre.

**Training Activity**   The centre has been conducting EDP courses for various States and Central government departments and international agencies. Over a period of time, the centre has trained a large number of officers in Electronic Data Processing. Besides, it has conducted six "Programmer" level courses and two "Training of Trainers" level courses in EDP for UN-sponsored candidates from the ESCAP region countries under the United Nations Household Survey Capability Programme between 1983 and 1991.

**Creation of National Data Warehouse**   The centre has initiated action for creation of a National Data Warehouse of Official Statistics (NDWOS). Under this project, the centre will preserve data generated by various Central and State government ministries and departments and public sector undertakings on electronic media, organise the data in the form of databases and provide remote access facilities to end-users through a network. The creation of the NDWOS will be completed in four phases. A large volume of data available with the centre has already been preserved on compact disks. These are now being disseminated to a large number of users.

# EMERGING PRIORITIES IN THE INDIAN STATISTICAL SYSTEM

## J. K. Ghosh[*]

## 1. Introduction

It was in the early years of independence that the Indian Statistical System began to change dramatically in response to the socio-economic needs of a newly independent, democratic country. The Reserve Bank of India (RBI), the organization of the census under the Registrar General and a number of other statistical establishments had already been created by the British. But the Central Statistical Organization (CSO), the single most important statistical organization since independence, was set up in 1951. The National Sample Survey (NSS), arguably the world's largest, had come up slightly earlier. Both were created to help remove gaps in the Indian database and played key rôles in decision making and in planning India's economic development. As we celebrate India's fiftieth year of independence, it is indeed the right time to take a look back as well as a look ahead.

This brief essay is organized as follows. Section 2 reviews some of the structural changes in Indian society and economy and their possible impact on the statistical system. Section 3 makes a few constructive proposals to solve emerging general as well as continuing problems. Section 4 concentrates on one particular area of priority, namely, health, and identifies a few concrete strategies for growth at a time of constraints on government expenditure. Section 5 contains concluding remarks.

## 2. The Changing Paradigm

Statistics in India has been closely linked to planning and massive state interventions. With the reduced importance of the state in capital formation, planning will have new goals. A very successful key player in the new scenario is the RBI. Such a reassessment of objectives and strategies is yet to be undertaken by the CSO or the NSS. The main activities of the CSO remain keeping the national accounts, measuring price changes and assessment of poverty levels. In each of these areas, one does see signs of progress. For example, the attention being given to small scale industries will improve national accounts, indicate strategies for generating employment and eventually eroding poverty. There is also a welcome sign of improvement in the database for India's environment. However, no major shift in focus, concept or structure is visible, subject to one qualification.

The Indian Statistical System is being modernized via computers in the matter of rapid collection of data and hopefully, rapid tabulation, analysis and communication.

On the debit side, it must be admitted that the Apex National Committee set up nearly a decade ago to strengthen the system has turned out to be singularly ineffectual. A former Director General of CSO has argued that the CSO should be de-linked from the Planning Commission and made part of a new ministry that will supply statistical information in various vital areas. It would certainly be of interest to study the Canadian, French and Dutch statistical systems from this point of view. How are they linked with the government? What provides their *raison d'être*? How is their autonomy protected? I have suggested a study of these three systems, because they have been favoured with gifted, innovative leaders and friendly, supportive government that understand the

---

[*] Jawaharlal Nehru Professor, Indian Statistical Institute, Calcutta

importance of statistical information. The International Statistical Institute may be able to help in such a study of the *rôle* of statistics in a changing world.

## 3. Directions of Growth

It seems to me that in the next century environment and health will be as important for us as, if not more important than, economic development. This is based on the assumption that the Indian economy will continue to grow at a reasonable pace, the old Hindu rate finally dying a natural death. The key *rôle* in managing the economy will pass to the RBI and other financial institutions. The national accounts will remain important from a long term point of view.

For short term fiscal policy, changes in national income have to be rapidly and accurately measured. I hope the CSO is aware of the need for closer links with the RBI.

On the other hand, as our economy improves, public policy of the government will be guided by considerations of protecting the environment or providing better health care to citizens who will be increasingly conscious of health issues in the face of a rising population and a degrading environment. I hope the CSO and the NSS will be key players in these two areas in addition to the statisticians in the ministries of environment and health. Since I will focus on health issues in the last section, let me turn to some other questions.

In addition to changing priorities, we are becoming aware of funding constraints. The bus is so full that a passenger in may mean a passenger out. I would suggest the CSO take a hard look at all its activities and see which remain relevant for decision making at a national level or measuring India's long term progress. Some statistics could sell well and be self-supporting- specially, financial statistics. Others should be supported if they continue to be relevant for measuring our welfare. Statistics that do not support these two tests are probably useless and need not be collected at all or collected only infrequently.

Statisticians within the system sometimes complain of continuing indifference of academics to data. Ironically, academics complain of the near impossibility of getting data, even at a price, from the NSS or the RBI. These are old problems assuming new importance and ought not to be difficult to solve. What seems to be needed is a time bound transparent policy of sharing data, possibly at a price, without breaking confidentiality agreements with respondents for information. One reason for the apathy of Indian academics is that hardly any in-depth study can be based on the summary tables that are regularly published. To attract the academics, they must be given access to raw data without impinging on confidentiality of sources. Modern computer based technologies can certainly ensure this.

Finally, cannot the Apex National Committee have a small functional subcommittee that can advise the CSO about new directions, new problems, new initiatives? Empirical studies in economic statistics in universities, hopefully, with collaboration from CSO, could provide the sort of analytical reports we had once expected from the CSO and NSS and which are regular features of strong statistical systems like the three that I have mentioned earlier. A notable exception has been a very striking paper by Professor B. S. Minhas comparing data supplied by the CSO and those supplied by the NSS. It had been possible to some extent because of his close links with the NSS at that time.

## 4. Our Ailing Health Statistics

In a democratic, developing country like India, whatever receives a lot of attention from the media tends to be taken care of by the government. Thus, as pointed out by Professor Amartya Sen,

post independence, India managed famines much better than China. In fact, since the mid-sixties, there has not been any famine in India.

In the health area, somewhat similar efficiency was shown in the eradication of malaria in the late sixties or early seventies and in the handling of plague a few years back. It is also true that malaria has made a come back but that is a different story.

As a sort of corollary to this, once again, as pointed out by Amartya Sen, whatever does not make headlines in newspapers tends to get neglected in India. Thus literacy, nutrition and primary health care remain neglected areas in India.

One sees the same neglect with respect to study or control of chronic diseases as opposed to rapidly spreading infectious diseases. One of the worst cases of apathy is the plight of some people in West Bengal who appear to suffer from chronic arsenic-poisoning. It is hard to believe, but it seems nothing at all has been done. On the contrary, there is even controversy about existing data.

I would like to focus on some general issues and what can be achieved by no more than a modest investment of resources.

The first task would be to identify the major chronic diseases in India, not only different types of cancer, heart disease, hypertension, diabetes and AIDS - all of which receive a lot of attention in the developed world - but also tropical diseases. One would also have to identify academic and Government institutions which work in these areas. Ideally, for each disease there will be a nodal centre where all available data and literature on the subject would be collected on a priority basis. For example, at the ISI, we try to subscribe to all important statistical journals and books as a sort of national resource centre. I envisage a similar voluntary *rôle* for identified institutions with respect to each disease. For example, at this moment the ISI has long term projects on cervical cancer, which is probably the most prevalent cancer among females in India. The ISI should then try to act as a resource centre for data and literature on this area and this activity should receive adequate funding for this both from the ISI and the appropriate State and Central Government agencies. Some of the institutions that immediately come to my mind are the ISI, the Indian Institute of Hygiene and Public Health, the Institute of Medical Statistics, the Department of Biostatistics at places like Vellore or the AIIMS. One can also include the Population Research Centres set up under the Ministry of Health provided the Ministry encourages health related activities also.

With respect to each disease, a statistical study would involve one or other of the followings aspects.

a) Database for mortality

b) Database for prevalence and incidence

c) Epidemiological and Intervention Studies

d) Database and literature on clinical studies

e) Sophisticated statistical "survival" analysis based on large epidemiological and intervention studies. "Survival" in this context means waiting time for whatever event one is waiting for

f) Stochastic modelling at a population or molecular level

I do not know of any large significant epidemiological or intervention studies in India. In contrast, such studies in developed countries and China receive worldwide attention. But in all other respects, i.e., in respect of all the other aspects that I have identified above, some work is taking place in India.

This lack of awareness and interaction between different groups working in India is all pervasive. At the present stage of our development one solution might be a newsletter. As the Internet becomes more easily accessible in India, an easier solution would be to set up a web page for each disease and work on it to contain all that is going on in India. Only with pooling of such resources major data and methodological bases will be possible.

It is a sad fact, of which I have become aware only recently, that even in respect of major diseases like cancer or diabetes, we do not have the database for calculating prevalence or incidence rates for either special risk groups like expecting mothers or the general population. Consulting physicians apparently have to depend on rates of other countries. Surely if we pool all the data that exist in India, something better can be done. Interesting methodological problems will have to be solved when data from different sources are combined. This would be a form of meta-analysis but not the usual kind where tests based on different clinical studies are combined.

Two other sources can be tapped. One is the computerized databases of private hospitals and medical insurance companies. Hopefully, these databases will provide linked records for the same patient over time without divulging a patient's identity. These can be the basis of longitudinal studies. The second source can be the NSS, which has experience in this area. If a subset of their field investigators is trained and provided with reliable automatic methods of determining the presence of disease, then we can get national and State level prevalence rates. A third source of data and methodology would be more interdisciplinary projects involving doctors and statisticians and joint projects between institutions. I have focused more on databases so far but statistics can also help in improving the quality of health care systems, disease management, better diagnosis and screening and identification of better drugs and treatments. All of these will require interdisciplinary effort. To sum up, further progress and coping with limited resources will need interaction and co-ordination and wide sharing and dissemination of whatever results are obtained. The newly established Survey Research and Data Analysis Centre at the Indian Statistical Institute can provide the *rôle* of a catalyst as well as a centre providing basic conceptual tools and paradigms.

## 5. Concluding Remarks

I have suggested a need for the CSO and NSS to examine their *rôles*, priorities and links with other government departments and the RBI in the context of changes taking place in India's economy and society. It seems to me that environment and health will be the emerging priorities in our statistical system. Hopefully, the CSO and the NSS will continue to be key players in these areas and will be as important in the next fifty years as they have been in the past five decades.

# THE NATIONAL STATISTICAL SYSTEM AND A CHANGING WORLD

S.P. Mukherjee[*]

## Introduction

Statistics (as a plural noun) have always provided a rational basis for objective decision-making by individuals and groups. Official statistics collected, collated and presented by State agencies are now recognized as indispensable inputs for planning, monitoring and evaluating public (or state) policies and programmes. In most of the countries, there exist national statistical systems charged with the responsibility (with or without the commensurate authority) to generate and disseminate adequate, reliable and timely data covering various facets of the economy, the society and the polity. Then there are the international statistical agencies (sometimes existing within the framework of some international organizations) to highlight inter-state comparisons as also to help formulate and evaluate global policies and actions.

Keuning (1997) points out that the main difficulty in studying interactions between economic development, environmental status and social change has been the lack of integration of statistics available on these aspects. If a data user tries to connect the multitude of statistical sources on his own, he soon finds that (a) figures on the same phenomenon differ by source, (b) often concepts, definitions and classifications do not match, (c) there are various holes in the data and (d) compilation of a consistent data set requires access to and know-how of concerned individual sources. While this lack of integration is much less in the case of National Accounts, there are many important social, environmental, demographic and related issues whose understanding and analyses suffer from this deficiency of available statistics.

The Indian Statistical System is one of the largest institutional frameworks, coming up with a vast wealth of information, not all of which gets due attention of decision-makers. It is one of the oldest also. With a large network supported by competent personnel and not-too-inadequate facilities for data management, the system claims a wide coverage of information items as well as of area and people. Of course, handicaps and deficiencies have not been unknown and one hopes that these will stabilize, if not disappear, with the passage of time.

In a world that is vastly changing very fast under the influences of globalisation and information revolution, the need for making efficient decisions on a whole array of vital issues in the face of uncertainty has been being strongly felt. Consequences of such decisions are not necessarily confined to the present or the near future nor to one particular country or nation. There has been a proliferation in the sources of data and new ways of data capturing have emerged. At the same time, there has been a growing appreciation of the inherent imprecision or fuzziness in social and economic measurements and observations, even when steps are taken to avoid biases or inaccuracies in the collection, processing and presentation of official statistics. Against such a backdrop, the Indian Statistical System merits a major overhaul.

[*] Centenary Professor of Statistics, University of Calcutta

**Need for an Adaptive System**

When we talk about a national statistical system, we can reasonably ask whether such a system possesses the four characteristics of a system (as mentioned in system science), viz. content, structure, communication and control. Such a system should be able to:

a. Identify its long-term and short-term objectives (sometimes differentiated as goals and objectives) as well as a strategy to achieve these objectives.

b. Ensure a structure in terms of job-differentiation by recognizing different types of information to be collected and presented and different aspects of information processing for which responsibility and authority should be clearly vested, possibly in different groups of people.

c. Establish clear lines of communication within itself to take care of inter-group or collective responsibilities as well as with the environment (comprising the government, the market and the society).

d. Review its content, structure and communication regarding their effectiveness (doing the right things) as well as efficiency (doing the things right) and modify or control one or more of these, as and when found necessary.

Coming to the first characteristic, one should appreciate that a statistical system has to be adaptive rather than teleological (with a fixed goal). In fact, the last characteristic 'control' refers to the ongoing arrangement for review of current goals and objectives, existing organizational structures and present practices for interaction along with the ability to change or modify any or all of these based on review findings.

To be adaptive, the national statistical system has to respond to emerging contemporary needs for information as well as contemporary practices for information management. It must appreciate that many new and unconventional sources of data like satellite-imageries have come to stay, many new concerns of policy makers like that for environment have to be attended to, many new techniques for data analysis like those that account for imprecision or fuzziness in data have since been evolved, and that the need for system effectiveness and efficiency has become imperative.

It has become imperative to develop (through a comprehensive dialogue) a statement containing a vision, a mission and a policy for the National Statistical System, to declare this (as a government pronouncement) and to disseminate the same (among the concerned functionaries). In this connection, the following statement by I. Fellegi (1996) during his 1995 Morris Hansen lecture is worth a note: "The objective of national statistical systems is to provide relevant, comprehensive, accurate and objective (politically untainted) statistical information. The end purposes to be supported by the information are multiple, but include prominently the monitoring of the evolution of the country's economic and social conditions, the planning and evaluation of government and private sector programs and investments, policy debates and advocacy, and the creation and maintenance of an informed public".

We may, as well, consider the vision statement of the Office for National Statistics (ONS) in the United Kingdom: An independent organization that inspires confidence and trust, ONS is to provide coherent, authoritative and trusted statistical information and registers, and to improve access to, and analysis and use of, its information. ONS exists to improve decision-making, stimulate research and inform debate within government and the wider community by providing a

quality statistical service; and to administer the registration of key life events. ONS is to serve:

  i. Government - by providing a service to support the formulation and monitoring of economic and social policies.
  ii. Parliament and the public - by informing them about the state of the nation and by providing a window on the work and performance of government.
  iii. Individuals - by registering key life events to help and protect them.
  iv. Business - by providing a service which promotes the efficient functioning of industry and commerce.
  v. Researchers, analysts and other customers - by providing a service which assists their work and studies.
  vi. European Union and other international users - by promoting international co-operation and comparability.

It is almost mandatory that the statistical system carries out planned and periodic audits of its own structure and function, as well as its goals and objectives. The much-vaunted National Advisory Board on Statistics was meant to effect these, but has been decimated to an apparently redundant and ineffective mechanism.

It will be a useful initiative to introduce a quality assurance system on the lines of the ISO 9001 standard in the establishments of the Central Statistical Organization (CSO) and the National Sample Survey Organization (NSSO) - the two key players in our national statistical system.

**Paradigm Shifts in National Accounts**

It has been argued (Griliches, 1994) that the share of the economy which is measured by official statistics with a degree of accuracy is declining. Environmentalists often mention about the gross natural product (as something not within the purview of GNP/GDP) that supports and sustains a sizeable chunk of our people remaining beyond the reach of contemporary development projects and measures. Between 1947 and 1990, the fraction of the economy for which productivity data can be deemed reasonably accurate fell from close to 50% to about 30%. While this is true for the world economy as a whole, the picture in our country is not radically different. In fact, current methodologies adopted for estimating GDP for the country or for a State have failed to reckon with what Goldfinger (1997) brands as fundamental changes (in the economy) - "incontrovertible, yet misunderstood". These are :

**a. The Changing Profile of Employment and Output Structure** While the numbers may vary from country to country, the over-all trend is definite : the share of industry and agriculture, whether in total output or in employment, has been falling steadily. Services represent the lion's share of both employment and output and constitute the principal - for many countries the only - source of employment growth.

**b. Globalisation** Foreign trade has been growing more rapidly than the world's output for decades. The World Bank estimates that world trade will continue to grow at roughly twice the rate of world output (World Bank, 1995). A key feature of globalisation is the increased mobility of people, driven by increased demand for international labour and leisure. This mobility has created a huge travel and tourism (T & T) business. According to an association of leading T & T firms, the World Travel and Tourism Council (WTTC), it is the world's largest industry. WTTC defines T & T as comprising five groups of activities: transportation, accommodation, catering/retail,

recreation/culture and travel related services. The WTTC has estimated that in 1991 T & T generated over 10% of the world's GDP (3400 billion dollars), employed over 10% of the global labour force (204 million people) and accounted for 11% of gross investment (WTTC, 1994).

   **c. The Ubiquity of Information Technology** (IT), which includes computers, telecommunications, and associated products and services (such as semi-conductors or software), is recognized as a structural vector that influences all economic activities. The development of IT has engendered a huge and rapidly growing new economic domain, whose global size is estimated at between 600 and 1100 billion dollars.

   And yet, while there is broad agreement on the existence of these three trends, there is no real consensus on their magnitudes, their underlying drivers, their structural causes and, more importantly, on their meaning. Actually, each trend has triggered fierce but rather inconclusive controversies. For mainstream economists and statisticians, and for many politicians, the shift to services is a puzzling occurrence that is inconsistent with their basic view of the economy. Finance is a quintessential service activity, yet it is excluded from international services trade data.

   In the intangible economy, a great deal of consumption and exchanges between economic agents either are free or are priced through mechanisms that don't reflect their full value to the consumer. In order to capture them, conventional measurement approaches, based on market transactions, have to be supplemented by metrics capturing the actual use of resources. For instance, estimates of household consumption should take into account not only monetary expenditures but also time budgets. This would allow a better capture of the consumption of entertainment artifacts such as a 'free' TV.

   Of late, there have been some attempts to extend the basic principles of national accounts to a wider range of statistics, namely, social and environmental statistics. A system called System of Social and Economic Accounting Matrices and Extensions (SESAME) has been designed by Keuning, elaborating on the general approach sketched in Chapter XX, Section C of the 1993 SNA (United Nations *et al.*, 1993). A precursor to SESAME is the SAM which is a matrix presentation of a sequence of monetary accounts, each showing a certain economic process and its relation to other economic processes. The alternative, double-entry book-keeping, serves well in presenting all transactions plus a balance sheet for a single agent or group of agents, but is less suitable for portraying inter-linkages among transactions of different types of agents. A SESAME breaks down money values into price (change) and volume (change). On the one hand, SESAME does not squeeze all welfare attributes into a single indicator. On the other, it integrates the meso-data and the concomitant macro-indicators, e.g. (un)employment, educational attainment and environmental degradation. Of course, the scope of the United Nations (1975) System of Social and Demographic Statistics (SSDS) is wider but its linkage with aggregate economic and environmental indicators is less explicit. Earlier, NAMEA was an exercise to incorporate contributions of various economic processes to environmental degradation with the traditional contributions to GDP, employment and foreign exchange earning.

## Some Emerging Needs/Uncharted Areas

   Data on some emerging areas like natural resource management, ecological changes, industrial metabolism, environmental impacts of urbanisation, dumping of foreign consumer goods, disposal of hazardous wastes, etc., are collected by a whole host of state or private agencies on a project mode. Such data are not always dependable, though their strategic worth is exploited by

proponents or opponents of a pre-conceived policy or a pre-meditated action. And even in situations where the integrity of such data is beyond doubt, concerned or potential users may not know of their existence or may be denied convenient access. It may, at the same time, be true that so-called official statistics do not exist on some such topic at micro-, or macro- or meso-levels.

Collection of relevant and reliable statistics on a phenomenon should keep pace with developments in our understanding of the phenomenon. For example, statistics to throw light on environmental pollution or ecological degradation are better obtained in terms of material flows through various anthropogenic processes like agriculture, industry, human settlement and associated services, etc. Despite the obvious linkages between chemical flows and the environment, industrial societies often possess little detailed knowledge of how major chemicals flow through the industrial economy after the production phase. They are, in fact, embodied in numerous products that are ultimately dissipated into the environment during use, disposed of in landfills or burned as refuse. In most cases, detailed quantitative information regarding the spread or accumulation of these chemicals in the environment is non-existent, and chemical monitoring is totally inadequate to fill this gap. What is urgently needed is a thorough book-keeping, that tracks the flows from mining, processing and manufacturing, consumer use and disposal. In fact, the recent concept of industrial metabolism provides a new vision for understanding how industrial societies produce, process, use, and dispose of materials and how these activities, taken as a whole, are linked to environmental changes.

Two things may be thought of in this connection. Firstly, some mechanism has to be evolved whereby agencies outside the ambit of the National Statistical System collecting data regularly or on an *ad hoc* basis may be compelled to follow accepted definitions and procedures. Secondly, such agencies may be required to provide a computer disc or a hard copy to an element of the Statistical System that may remain responsible to collate and classify all such data sets after a proper check for their usefulness and usability. This element should also arrange for the dissemination of such data sets to those who want to process or analyse these. To start with, these two stipulations may be implemented in respect of state-sponsored projects requiring collection of data at a macro-level (e.g. district or a bigger geographical area, an industry group, etc.).

All this may require framing and implementing an appropriate legislation. What is more important, however, is an exercise (maybe by the official statistical system itself) to propagate the message of quality in statistics, to promote the use of statistics and to deprecate abuses and misuses of statistics.

While on the topic, it may be pointed out that statistical meta-analysis (Hodges and Olkin, 1985) can be very well applied for the purpose of combining data or information from different independent sources which are targeted at a common goal. The task of sourcing isolated studies by individuals and groups in government, industry, trade, academia, national laboratories and kindred organizations and synthesizing or integrating their summary findings through statistical methods may be taken up by the National Statistical System to develop workable databases on some emerging topics not currently covered by official statistics.

## Uncertainty

Wherever estimates of aggregates or averages or proportions are offered by the National Statistical System, it is understood (by the producer, if not fully appreciated by the users) that such figures are based on a sample (not necessarily random) and that these are also affected by non-sampling errors like measurement inaccuracies, response biases or recording errors. As a result, any

such figure should be accepted only along with an estimate of the uncertainty associated with the figure (or measure/count). The standard error of the estimate is just one component of the overall uncertainty that includes systematic components. Environmental measurements on a macro-scale are often derived as the product of a micro-level figure multiplied by a very large scaling-up factor. And the micro-level figure itself might have ignored systematic variations like seasonal or diurnal variations. Inherent imprecision of measuring instruments or of measurement method also contributes to uncertainty of measurements relating to stocks and flows, concentrations and budgets of pollutants, for example.

In this age of globalization, the need for international comparability of economic indicators is strongly felt. The accuracy of such comparisons depends partly on the accuracy of the national data concerned and partly on the accuracy of methods for such comparisons. Computations of such aggregate measures as Human Development Index, Human Poverty Index, Income Gap Ratio, etc., are subject to a whole host of assumptions like equitable distribution or rational consumption or identical needs. Comparisons of such measures over a long period of time within the same country or at the same time across various countries are beset with genuine problems of index number construction as also of social and cultural transitions. Widespread use of computers has, to some extent, reduced uncertainties in reported figures, but the same cannot compensate for genuine uncertainties in measurements and observations. Not too unoften, policy-makers read too much into differences in decimal points given for GDP figures or growth rates. It will be wise on the part of a National Statistical System to work out and declare limits of uncertainty in such figures.

**Concluding Remarks**

Before concluding, let me quote Goldfinger (1997) : "Today comparability becomes more laborious to achieve as goal posts are being continuously moved. More importantly, users want data that signal future trend and performance and help them to cope with rapid and unexpected change. Statistics need to become more and more like meteorology. Their main value should reside in the ability to provide early warning signals and inflection points indicating shifts in major trends. The traditional rôle of statistics was to create a map of the present. They now have to provide markers for the trajectories to the future".

India had a head-start in statistics over many other countries - both developed and developing. India does not lack in expertise or in understanding, though lack of political will to cash in on a comprehensive statistical database for decision-making has not been non-existent. All that we need is to integrate and consolidate the available expertise and resources and transmute the same into leadership to initiate concerted action. Let us hope that this is realized during the historic year that marks the 50th anniversary of India's Independence.

---

**References**

1. Ayres, R.V. (1978). *Resources Environment and Economics: Applications of the Materials/Energy Balance Principle.* John Wiley, New York.
2. Fellegi, I. (1996). "Characteristics of an effective statistical system", *International Statistical Review, 64,* 165-197.
3. Goldfinger, C. (1997). "Intangible economy and its implications for statistics and statisticians", *International Statistical Review, 65,* 191-220.

4. Griliches, Z. (1994). "Productivity, R & D and the data constraint", *American Economic Review, 84.*

5. Hodges and Olkin (1985). *Statistical Methods for Meta Analysis.* Academic Press, New York.

6. Kennessey, Z. (1997). "A perspective on the accuracy of economic observations", *International Statistical Review, 65,* 247-259.

7. Keuning, S.J. (1997). "SESAME : an integrated economic and social accounting system", *International Statistical Review, 65,* 111-121.

8. Morgenstern, O. (1963). *On the Accuracy of Economic Observations.* Princeton University Press, Princeton.

9. Mukherjee, S.P. (1994). "Quality - A systems approach", *European Quality.*

10. Office of National Statistics (1996). *Vision and Values.* London Office for National Statistics.

11. United Nations (1975), Towards a system of social & demographic statistics, Studies in Methods, Series F, No. 18, New York.

12. United Nations (1993), *Integrated Environmental & Economic Accounting,* Studies in Methods, Series F, No. 61, New York.

# NEED FOR A RELIABLE DEMOGRAPHIC DATABASE IN INDIA

## K.B. Pathak[*]

The utility of up-to-date and accurate demographic database hardly needs to be emphasized for the purposes of development planning. While decennial censuses conducted in India provide fairly adequate data on the population stock over time, data on the main population flows, namely, births, deaths and migration in a given period of time remain deficient and incomplete. Since 1970 onward the Sample Registration Scheme (SRS) has been able to provide comparatively good estimates of vital rates at the State and national levels for the rural and urban areas separately. Specific demographic surveys, on the other hand, have been conducted to throw light on the causes and consequences of changes in demographic processes over time, not only at the State and national levels but also at the levels of smaller areas.

In fact, demographic sample surveys covering larger areas have proved rather alternative to the Civil Registration System for providing estimates of the vital rates needed in evaluation of the on-going family planning programme in the country. Recently, National Family Health Survey (NFHS), the largest ever national representative sample survey covering 89,777 ever-married women aged 13-49 both in rural and urban areas provided a treasure-trove of information on fertility, family planning, child mortality, maternal and child health and immunization of children. The results of the survey are contained in 20 detailed reports giving the demographic and health scenario for 25 States and the nation.

The results of this survey have been widely disseminated and used in formulating population policies and programmes at the national as well as the State level. Since in recent years the focus has been on decentralization of planning up to the level of district and block, it seems that we have to increase sample size and the coverage of various surveys to provide estimates at the district level. While the idea of conducting national-level comprehensive surveys is laudable to provide checks on the regular system of data flows, it must not be forgotten that repetition of such surveys before the lapse of 5 years is not advisable because of the time and cost involved. Moreover, analysis and utilization of data take some time. Much of the census data of 1991 is still not published for wider use. Collection of data through official systems has not only been found deficient but also has not been made available in time when most required for planning.

There is a need to strengthen the system on a regular basis. Already efforts are being made to improve the registration of births, deaths and marriages. Provision of electronic equipment at the grass-root level will certainly improve the collection and compilation of data as well as their onward transmission to higher levels. It will also facilitate the dissemination of information very fast at the district or even block level. Already the idea is being mooted to dispense with routine census-taking. However, we must ensure that a parallel system which is less cumbersome and less time-consuming should be evolved before we do away with the population census. In case civil registration is perfect, hundred per cent household surveys with less cost can be organised throughout the country with the help of health functionaries spread over the whole country in just two to three months with almost 100% accuracy to give the total number of households, their residents along with their ages and some basic characteristics. These household surveys will enable the health workers to assess

[*] Director, International Institute for Population Sciences, Deonar, Mumbai.

also births, deaths and new marriages annually in the area. This is not at all difficult; in fact, in some States like Tamil Nadu, Madhya Pradesh and Rajasthan, such surveys have been conducted. All the eligible couples have been interviewed. Similarly, persons engaged in development and agriculture departments can be asked to conduct household inquiries on employment, occupation and persons engaged in organised activities. We have a large number of skilled and semi-skilled government servants in Central, State and local governments and the canvassing of schedules in the selected sectors will not be a difficult work. What is more required is to make them aware about the importance of data in effective provision of their services and encouragement for them to better equip them with skill for handling socio-economic data with a view to future prospect of their career improvement. This is a question worth pondering: how some non-governmental organisations are able to do better in collecting demographic data than the governmental agencies engaged in such jobs?

Strengthening of the official data system in the field of population, however, does not negate the need of the specific demographic surveys to study in detail the demographic behaviour of the population either at the national or at the sub-national level. However, development of new techniques in different fields of population studies can effectively help in bridging the data gaps. It is here that statistician demographers have to play their *rôle*. They must come out with better and improved techniques which do not depend on extraneous data needs and assumptions and rather depend on use of the indigenous data already available. Of course, this requires continuous research efforts, especially at the level of some well-known research organisations.

**Estimation of Population Parameters**

In any sample survey, our aim is to provide reasonable and acceptable estimates of various population characteristics based on the sample observations. These estimates are generally expressed in terms of averages and proportions, e.g. average number of children born/alive/ desired/ideal, birth rate, death rate, infant mortality rate, etc. When the sample design is as simple as simple random sampling, the sample average or proportion can be safely used to obtain the estimates of the population parameters. In the complex sampling designs as adopted in various large-scale sample surveys like the NFHS where the units are considerably heterogeneous, the direct use of sample averages or proportions will provide considerably biased and misleading estimates.

In a large-scale sample survey with a complex sampling design, one has to give proper weightage to the sample estimates of various sections (or strata) so as to obtain acceptable estimates of the population characteristics for the whole domain of study.

A good deal of discussion regarding the use of appropriate sample weights can be found in the State-level reports of the NFHS. The interesting feature of the NFHS weights is that they provide adjustment factors (weights) for non-response in contrast to many other large-scale sample surveys conducted in India.

For the purpose of fruitful planning and policy-making pertaining to health and family welfare, it is necessary to have basic information at smaller regions within a State, such as districts or groups of districts. These data will be useful to chalk out requisite action programmes for the areas under study. For obtaining valid estimates of these smaller sub-divisions, it is necessary to increase the over-all sample size which will definitely increase the work load and in turn may increase the non-sampling errors as mentioned earlier. Necessary precautions should be taken to minimize the non-sampling errors as much as possible either by reducing work load or by increasing

field supervision during the data- collection and data-processing phases through the use of skilled and well-trained personnel.

In China, a one-in-thousand post-census sample survey was conducted in 1982 and it was found sufficiently representative to give estimates of important social, economic and demographic indicators even at country and State levels. In case of India, it would mean covering over 1.80 lakh households, almost double the number of households covered in the National Family Health Survey (NFHS). Further, every year, one-third of these households can be replaced to account for changes occurring in the population. Virtually, this much sample size scattered all over the country might be able to give estimates of major social, economic and demographic indicators even at the district level. As a matter of fact, some innovative survey technology has to be evolved to reduce data gap existing at the district level, on the one hand, and to standardize the data set, on the other hand. Probably, the sample size at the all-India level has to be almost triple of 1.80 lakh households to encompass both SRS and NSS coverages. If there is proper counseling, all data-collecting agencies at the national and State levels can pool their resources and streamline their efforts to collect relevant and most needed data by canvassing compact and less time-consuming shorter schedules. Even census-taking can be considered for only 10 per cent of the households and thereby time in conducting and processing the relevant data can be reduced considerably. There is an urgent need to reconsider new approaches to conducting population censuses and related surveys.

# WHITHER STATISTICS ?

## J. Roy[*]

## 1. Introduction

The fiftieth anniversary of our independence is a right occasion for us, who belong to the statistical profession as teachers, researchers or practitioners, to look back and examine the record of our service to the nation, note our successes and failures, and then look ahead and ponder how we can improve our performance in the future. The main concern in this account will be what is known as the official statistical system of India. We shall take a quick look at early developments under Prasanta Chandra Mahalanobis (PCM), the current situation, then attempt a critical assessment of successes and failures, and finally offer a few suggestions.

## 2. Early Developments

Collection and use of statistics for administrative purposes in India has a very long and well-known history, spread over many centuries. However, it was only in the nineteen thirties that Prasanta Chandra Mahalanobis (PCM) put statistical practice on a scientific basis in India. The other pioneer was Pandurang V. Sukhatme (Sukhatme), who specialised in application of statistical methods in agriculture-related subjects and nutrition.

PCM demonstrated that statistics is an essential tool for rational decision-making based on fallible information in any activity - science and industry, business and trade, government and administration. He called statistics *the key technology* of the age.

With fanatical zeal, PCM pursued the development of statistics in India in all directions. His own theoretical research covered many areas of statistics - sample surveys, multivariate analysis, developmental planning - to mention a few. In each of these areas, he made path-breaking advances. Most of his research was motivated by logical and operational problems of collection and interpretation of fallible data and substantiated through large-scale applications.

He established the Indian Statistical Institute (ISI) in 1931 as a scientific society for theoretical and applied research in statistics. A brilliant group of researchers - Raj Chandra Bose, Samarendra Nath Roy, C Radhakrishna Rao, K. R. Nair, Kanwar Kishen and many others joined him at the ISI. *Sankhyā*, the research journal of ISI, soon won international recognition. It is interesting to note that K. R. Nair brought his academic excellence into the official statistical system of India first as the Senior Statistician of the Forest Research Institute and then as the Director of the Central Statistical Organisation. So did K. Kishen as the Director of the Statistical Bureau of Uttar Pradesh. A strong relation between academicians and practitioners of statistics was thus initiated.

Even in the midst of the struggle for independence, the Indian National Congress had started thinking about strategies for economic development of the country after independence. It was at that time that PCM offered Pandit Jawaharlal Nehru his full assistance in all statistical aspects of developmental planning. Thus began an association which lasted till Panditji's death. As soon as

[*] Research Professor Emeritus, Indian Statistical Institute, Calcutta and former Member of the Governing Council, NSSO

Panditji took over as the first Prime Minister of independent India, he invited PCM to be the Statistical Adviser to his Cabinet.

PCM utilised this position of authority to make statistics serve a key *rôle* in planning for economic development of India. The most important single indicator of the economic level of a country is its national income. Current estimates of this parameter were not available. A National Income Committee was set up with PCM as Chairman to remedy this defect.

To cover gaps in information needed for calculation of national income, PCM created the National Sample Survey, the largest and arguably the most scientifically designed repetitive large-scale multipurpose sample survey in the world.

PCM was largely responsible for consolidating the Central Statistical Organisation (CSO) and the Department of Statistics (DoS) to put the official statistical system of India into an organised structure.

He fervently believed that the discipline of statistics has an important *rôle* to play in socio-economic development of our country. He demonstrated this by working out a draft plan frame for the second five-year Plan. He thereby showed how allocation of scarce resources between competing projects can be done on a rational basis.

Earlier, PCM had started at ISI training and certification programmes for professional statisticians. In the academic area, he began an M.Sc. degree course in Statistics at the Calcutta University from 1941 and a Bachelor's Honours degree course in Statistics at the Presidency College, Calcutta from 1944. Later, when ISI was recognised as an Institution of National Importance by the government of India, it also started awarding degrees in Statistics. The DoS also developed programmes for in-service training of statistical officers.

Another pioneering effort of PCM was to introduce electronic computers in statistical work in India. The first ever electronic computer in India - a Hollerith HEC-2M system from UK - was installed by PCM at ISI in 1949. The DoS now makes extensive use of its own computers in its work.

PCM also introduced Statistical Quality Control methods and Operations Research techniques for improvement of performance of industries in India. The effort continues at ISI, but strangely, the DoS has kept itself out of this kind of endeavour.

### 3. The Current Situation

Consistent with the federal structure of India, our statistical system is also decentralised. Basic data are collected by various departments of Central and State governments, the Reserve Bank of India and other financial institutions, local bodies and many organisations in the private sector. Each State has its own Bureau of Economics and Statistics, though are called by different names in different States.

The Department of Statistics (DoS) in the Ministry of Planning and Programme Implementation of the Central government is the apex body in the statistical system of India. It is responsible for initiating, supporting, standardising and coordinating major statistical projects. The DoS has three major wings - the Central Statistical Organisation (CSO), the National Sample Survey Organisation (NSSO) and the departmental Computer Centre (CC). In addition, the DoS serves as the administrative link between the ISI and the Government of India.

A distinct feature of the statistical system of India is the Indian Statistical Service (ISS) - a professional cadre of trained statisticians who man important statistical positions in all government

departments. This was set up in 1961 by pooling together statistical personnel from different government departments. Since then, new recruitments have been made on the basis of stringent competitive tests by the Union Public Service Commission.

The CSO maintains liaison with Central, State and international statistical agencies. On its own, it prepares national accounts, collects and publishes industrial statistics, conducts economic censuses and follow-up surveys, middle class family income-expenditure and other important surveys. It also conducts training programmes for ISS probationers and other statistical personnel.

Population census and agricultural statistics are two major areas where the CSO plays only an external consultative *rôle*. The decennial population census of India has a very long history. For administrative and other reasons, it has all along been conducted by the Census Commissioner, a very senior and scholarly administrator in the ministry of home affairs. There were some attempts in the fifties to bring the census of India under the control of the DoS, but this was overruled by the government.

Likewise, agricultural statistics has been by and large the responsibility of the ministry of agriculture. In pre-independence India, PCM had conducted a number of large-scale sample surveys to estimate the area and yield of jute in Bengal and of wheat in Uttar Pradesh, using a method *(small circular cuts)* substantially different from that developed by Sukhatme *(large rectangular cuts)* for the ministry of agriculture. During the fifties, when PCM was the statistical adviser to the cabinet, there was a long and interesting proxy debate between PCM and Sukhatme about the appropriate methodology for estimation of area and yield of crops. After a number of critical comparative studies, compromises were made which led to significant improvement in methodology.

The NSS was created in 1950 at the initiative of PCM and with whole-hearted support from Panditji. The objective of the NSS is to carry out nation-wide annual sample surveys to cover the gaps in data required for estimation of national income and for preparation of development plans. Initially for nearly two decades, the ISI was responsible for survey design, data processing and report writing. A field organisation was set up under the Central government for collection of data. Later, an organisation (NSSO) was created in the DoS to take up complete responsibility for the NSS under the technical control of an autonomous Governing Council to which ISI nominates two members. The autonomous Governing Council is meant to ensure that the data collected are free of administrative bias.

Since its inception, the NSS has made significant additions to the national database, particularly in the following areas: (i) household consumption - leading to many studies about the extent of poverty in the country, (ii) employment and unemployment - sorting out conceptual difficulties in the definition of employment in the Indian context, (iii) manufacturing and trade - in the unorganised sector, (iv) land holding, (v) crop area and yield, (vi) literacy and education, (vii) morbidity, disability, maternity and child care, (viii) utilisation of public distribution system, (ix) housing condition, (x) prices, and (xi) household debt and investment.

A departmental Computer Centre (CC) was set up in 1967 to serve the data processing needs of the Central government. When the Department of Electronics (DoE) established the National Informatics Centre (NIC) with its own network of computers (NICNET), much of the work was transferred to it. The major responsibility of the CC has since then been to process data collected through the NSS and surveys conducted by the CSO. In the past, processing of NSS data at the ISI on Unit Record Machines and later at the CC on computers used to take a much longer time than anticipated, primarily because of the difficulty in removing internal inconsistencies from the fallible data collected from the field - a process metaphorically but aptly called *cleaning*. Data cleaning is essentially a statistical procedure. Only competent statisticians, experienced in the field of

application, are capable of developing an efficient data-cleaning procedure. The procedure can then be computerised. If this essential pre-requisite is not fulfilled, even the largest electronic computers with the most brilliant computer software engineers would fail to produce results in time. This elementary lesson was learnt at a great cost when the NIC was entrusted with the processing of an entire round of NSS data with disastrous consequences. The summarised data never saw the light of the day. The NIC has been transferred from the DoE to the Planning Commission.

٭       The importance of data cleaning and the *rôle* of concerned statisticians in developing appropriate procedures are better understood now. Statisticians in the NSSO are now themselves engaged in processing of NSS data on in-house computers at the NSSO. Though these computers are only of moderate power, there has been noticeable improvement in cutting down delays in processing of data. This is indeed very heartening.

## 4. Achievements

At the time of independence, India inherited a statistical system in the shape of independent statistical units attached to different ministries and offices, each using its own concepts and definitions and procedures, relying by and large on common sense and intuition for interpretation of collected data. Data were obtained primarily as a by-product of administration and in a few cases through periodical complete enumeration. There was little or no assessment of the reliability of data, and little or no attempt to make logical use - as distinct from intuitive use - of statistics in decision making.

Since independence, there has been tremendous progress in all directions overcoming many difficulties. Under our federal structure, statistics is the concern of the States as well as of the Central government. Allocation of resources from the Central government to the States depends on statistics reflecting socio-economic conditions and performance of the States. It is not unlikely that statistics collected by individual States may be biased in favour of the respective States. It is absolutely essential at the Central level to standardise concepts, definitions, and procedures of collection of statistics so as to get rid of this likely bias. The CSO was created mainly for this purpose and it has been largely successful in this area.

The DoS through the NSS has successfully implemented the methodology of collection of data from statistically designed large-scale sample surveys.

Use of statistics in the construction of economic models for developmental planning, first demonstrated by PCM, has now become a standard practice. The CSO has helped the Planning Commission in a very significant way in this area.

Creation of the ISS has made practice of statistics an important career in government service. Statistical education has spread far and wide. Courses at the Post and Undergraduate levels are now offered by most Universities in India. Statistics is also taught at the school level. Professional training in various areas of applied statistics - agriculture, industry, commerce, data processing etc. - are provided by various organisations of repute.

## 5. Weaknesses

There are three major aspects of the quality of a body of statistical data: their *relevance* in respect of the phenomena being studied, their *reliability* as measured by the extent of sampling and non-sampling errors present in the data, and their *timely availability.*

An astonishingly large volume of data are collected all over the world - and India is no exception - as a ritual, as a matter of routine, with no use in view. It is because of this that PCM

emphasised again and again that statistics *must have a purpose.* The NSSO, however, deserves praise for the thought and care it gives to decide upon what information is to be collected in a particular round of the survey.

The main contention of the discipline of statistics is that *all* empirical information, *all - without any exception whatsoever,* is contaminated with some amount of error. The amount of error, the opposite of which is *accuracy,* can be controlled, but can never be entirely eliminated. Fortunately however, it is not necessary for the data to be absolutely error-free. Some amount of error can always be tolerated. The margin of tolerable error generally depends on the purpose for which the data are to be used. It is therefore an ethical necessity for the *producer* of statistics to provide to the *user,* an estimate, however crude, of the margin of error of the statistics. A large failure of the DoS has been that it has seldom provided to the user any idea about the margin of error. It is regrettable that data provided by the DoS have to be taken on faith - not on the basis of any assessment of the margin of error.

Another failure has been the enormous delays in publishing results, particularly in the case of large-scale sample surveys like the NSS. However, as already noted, of late there has been considerable improvement in this area.

The DoS has failed to incorporate into practice much of the methodological developments in statistics that took place during the last 40 years. Except for minor adjustments, the basic design of the NSS is still the same as it was when the NSS was started. It is not very well known that the original sampling design for the NSS - *two-stage stratified sampling in an interpenetrating network of sub-samples, with probability of inclusion of a unit proportional to size at the first-stage and uniform at the second, the actual selection being done in a circular systematic manner* - was actually a plan worked out by the renowned statistician Debabrata Lahiri at the ISI to meet a contingency. Until then, all large-scale surveys conducted by ISI used the technique of *areal sampling* for selection of units for investigation. Large-scale cadastral survey maps were needed for this. In the short time available for selecting the first-stage units to be surveyed in the very first round, it was just not possible to collect all the maps from all over India. It was necessary to devise a sampling plan using lists of first stage-units instead of maps. That the design adopted turned out to be quite an effective design is another matter. The efficiency of any two-stage sampling design depends on the number of first-stage units to be selected from a stratum, and the number of second-stage units to be chosen from each selected first-stage unit. The statistical theory is well known, but no serious attempt has been made to use the theory in the context of the NSS. It is a matter of shame that, even after fifty years of operation, we have not even an approximate idea about how the sampling error depends on the number of sampling units at the various stages of sampling.

Though powerful techniques exist by which auxiliary information can be exploited to improve the accuracy of estimates obtained from current surveys, no use has ever been made of such techniques. For example, in the case of repetitive surveys like the NSS, information from previous rounds, or from neighbouring areas, or from external sources can be used to improve the accuracy of estimates for the current round, through time series analysis or regression techniques. But this has never been tried.

There is a great demand for information at the grass-roots level. No attempt has, however, been made for applying *small-area estimation techniques* that enable one to get fairly reliable estimates for small areas from nation-wide sample surveys without unduly increasing the size of the sample.

Likewise, no attempt has been made by the DoS to make use of Remote Sensing Technology, for the purpose of improving estimates for areas under different crops, or for study of

changes in habitation pattern, or for that matter for study of land use. A modest beginning has been made by the Indian Council of Agricultural Research, but the DoS, which is much better equipped, is yet to enter into the picture.

The quality of theoretical research in Statistics in India has been highly acclaimed internationally. Unfortunately, this is not reflected in the quality of professional statistical work. Most technical problems faced in statistical offices in India today are tackled intuitively using a "common sense" approach. Very rarely is any use made of research results obtained elsewhere, or an independent research investigation undertaken to solve the problem.

Computers have been introduced in statistical work but merely for data processing. No use has yet been made of computers for design and control of statistical operations.

A singular failure of the DoS has been that it has been unable to create in the public mind a fair picture of its activities - the invaluable service it provides to the nation. In the absence of careful explanation, many of its announcements lack credibility in the public mind. The announcements on growth rates of gross domestic product, and their frequent revision, create confusion in the public mind - to say the least. The low inflation rates that are announced from time to time are totally in conflict with the high rise in prices that the consumer has to face every day in the market place. No attempt has been made to explain to the layman, in simple language, the basic concepts behind the various statistics periodically announced by the DoS and how these should be interpreted.

## 6. Looking Forward

Though the DoS has made significant progress in making statistics an important tool for administration and developmental planning, there is of course always a scope for further improvement. I enumerate below some points which deserve the attention of the authorities:

**Motivating Statisticians**  Statisticians in Central and State Governments are expected to introduce innovative techniques for collection, analysis, interpretation and dissemination of statistics and not merely to be content with routine management of such activities. They should themselves take up critical studies on deficiencies of current statistical practice and develop remedial measures. Their job should not end with publication of statistics, but beginning from there, they should also explain their policy implications. The DoS should arrange frequent training programmes for statistical officers on new methodologies in sample surveys, time-series analysis and forecasting, computer applications and other relevant topics - in the department, or in research institutes in India or abroad. A mechanism should be established for quick recognition of statistical work of merit. A point of importance is that training in statistical computing is much wider and of greater relevance for professional statisticians than a course in mere computer software engineering.

**Motivating the User**  The DoS should bring out regular publications at a popular level to bring to the attention of all users of data the following matters: (i) the structure of the Indian Statistical System, (ii) current data availability, (iii) current controversies relating to quality of statistics and (iv) methodology for computing specific indices and proper interpretation of the indices. The DoS should lay down a procedure for making raw data available to researchers.

**Simplification of Survey Design**  NSS schedules are too long and complex. These should be simplified and/or split so that an informant can supply the information in, say, thirty minutes. To increase the size of the Central sample to get reliable estimates for "small areas" is not practical. Small-area estimates should be obtained by States through decentralised methods. Sophisticated new techniques of small-area estimation should be tried on data from Central samples to provide a

background for comparison. Extensive sampling experiments should be carried out regularly to improve the overall efficiency of sampling design.

**Computerisation** Computers should be used more vigorously not only for processing of data, but also for sampling design, control of statistical operations and dissemination of information - through printed reports, floppy-disks, compact disks and Internet. A statistical communication network should also be set up and Internet facilities should be introduced in all important statistical offices.

**Academic Training** Present statistics curricula at both the Bachelor's and Master's levels are not adequate for official statisticians. Greater emphasis is needed on application of statistics in planning and administration, statistical operations, sample surveys, time-series analysis, statistical modelling and use of computers. As one of the largest employers of university students of statistics, the DoS should immediately take up with the University Grants Commission and the Human Resources Development ministry the need for an urgent revision of the academic programme in Statistics. The selection test for recruitment to the ISS should put greater emphasis on the above practical topics than on abstract theoretical aspects.

**Applied Research** The university system as it exists today does not provide much incentive for applied research in statistics. The main reasons are: (i) except for a few notable individuals, most college and university teachers are not in contact with the DoS or other agencies involved in professional statistical work, (ii) a teacher or a research scholar in a university has very limited access to official statistics, and (iii) computing facilities in most universities are usually very poor. It is therefore necessary for the DoS in its own interest to step in and sponsor or support applied research in statistics in areas that are important to the DoS. An illustrative, but by no means exhaustive, list of areas where applied research is urgently necessary is given below:

i) Sampling design for multipurpose large-scale surveys
ii) Estimation of sampling and non-sampling errors
iii) Use of auxiliary information from past surveys or other sources to improve current estimates
iv) Logical basis for imputation of missing observations
v) Generalised software for survey data processing
vi) Management and control of large-scale surveys to avoid time and cost over-runs
vii) Small-area estimation
viii) Use of Remote Sensing Technology in surveys
ix) Identifying the poor
x) Short-time forecasting
xi) Modelling the Indian economy, etc.

It is very heartening to know that the DoS has already taken up a project specifically for this purpose.

# SOME SUGGESTIONS FOR STRENGTHENING THE CURRENT NATIONAL STATISTICAL SYSTEM

## K.C. Seal[*]

,       After independence, great emphasis was laid on restructuring the then statistical system to ensure collection of data pertaining to various sectors of the economy in a cost-effective manner. These data were badly needed for drawing up an effective development plan for the entire country. The short background paper prepared by the Department of Statistics (DoS) brings out clearly the comprehensive nature of the current statistical activities of the DoS. This was possible primarily due to the relentless efforts of Professor P.C. Mahalanobis (who served as Honorary Statistical Adviser to the Cabinet till his death) and the technical support provided by the Indian Statistical Institute (ISI) in the 50's and 60's. The National Sample Surveys designed by ISI in the 1950s served as a model for the United Nations Statistical Office, New York (UNSO) for drawing up a new project called "UN National Household Survey Capability Programme" in 1977 under which a large number of developing countries were assisted by the UNSO in generating minimum data pertaining to various sectors of their economy as required by the respective countries.

After the death of Professor Mahalanobis, the post of Honorary Statistical Adviser was not filled up. This seems to have resulted in assigning much lower priority to new statistical schemes by the Central Government (i.e. the Planning Commission and Ministry of Finance). Consequently, not much improvement in the National Statistical System could be registered. A high-level Committee was set up by the DoS in 1979 to review the then National Statistical System and this Committee made a large number of recommendations [1] for strengthening the National Statistical System. But only a small proportion of the recommendations could be implemented so far, due to various constraints. The changing scenario of quality of official statistics and statisticians, after Professor Mahalanobis, is described briefly in [2]. In recent years, there has been a growing concern amongst many professional statisticians and economists about the deteriorating quality of official statistics which are being disseminated by the government departments [2,3,4]. Although weaknesses in the system and remedial measures have been critically examined by several researchers [2,3,4,5,6,7], no effective action plan has been evolved by the DoS so far to improve the quality of official statistics and statisticians after a careful review of the various suggestions which were made earlier.

In the following paragraphs, areas of major concern pertaining to the Central Department of Statistics are highlighted. Specific suggestions for improvement of the prevalent system are also briefly indicated.

## Central Statistical Organisation (CSO)

The timeliness of many publications needs considerable improvement. Although large-scale computerisation has been introduced, the present situation leaves much to be desired. Some of the useful publications of the CSO which were regularly brought out and widely appreciated in the past by the public seem to have been discontinued by the CSO presumably on grounds of economy.

The standard-setting *rôle* of the CSO has not been performed satisfactorily. Technical manuals on newly-emerging topics of interest suggesting detailed practical steps to be adopted for generating meaningful data on these new issues, such as study of the extent of empowerment of women, degradation/pollution of environment and deprivation of basic human needs amongst the people, would be welcomed by the general public.

---
[*] Formerly Director General, Central Statistical Organisation, Government of India.

Although considerable efforts are made to improve the timeliness of CSO's two important publications pertaining to National Accounts Statistics (NAS) and Industrial Statistics relating to the registered sector, real improvement in timely dissemination of "quality statistics" to the public has not taken place so far. While efforts are being made to produce many new tables in the National Accounts Statistics (NAS) document keeping in view the recommendations of the UNSO by utilising whatever fragmentary data are available in the country from various sources like follow-up surveys of economic censuses and comprehensive cross-sectional agro-socio-economic data collected on a regular basis by the NSSO over a long period, no serious efforts have been made to study the "margin of uncertainty" (the combined effect of sampling and non-sampling errors) of the estimates of at least the important parameters published in the NAS document for the benefit of data-users.

The primary *rôle* of the CSO as the nodal agency in coordinating various sample surveys carried out by diverse government and semi-government agencies has not been played very effectively in avoiding the divergent conclusions drawn by the CSO and some of the administrative departments, particularly those relating to agriculture and a few other sectors which are generally considered State subjects under our Constitution. The quality of published statistics pertaining to these sectors is seldom critically reviewed by a competent statistician at the highest possible level. A need for high-level posts of Statistical Advisers in these Central government departments is keenly felt by non-government researchers and professional statisticians. As an alternative, a new post of a high-level Statistical Adviser attached either to the Prime Minister or the President of India with effective power to review and guide the major statistical activities in all the sectors of the economy and ensure prompt dissemination of important results should be created and filled up by an eminent scientist. Statistics should be treated as a part of "Information Science" rather than for planning and evolution as viewed at present [1]. Results of statistical analyses should be widely disseminated to the general public in an impartial manner. This will help the policy-makers to take correct decisions on various issues which are considered relevant.

## National Sample Survey Organisation (NSSO)

This is an organisation which has gained world-wide recognition for the innovative approach to generating a comprehensive database within a short period for the entire country with the help of well-trained field investigators and supervisors. An over-view of the current position is available in [8]. Two major shortcomings of data obtained through the National Sample Surveys which are being conducted at present on a matching basis by the NSSO at the Centre and State Statistical Bureaus (SSBs) at the State level are yet to be overcome.

Pooling of the sub-sample estimates based on both the Central and the matching State sub-samples for the same socio-economic characteristic would normally provide a more dependable estimate and also give an idea of the "margin of uncertainty" of these estimates. Although sub-samplewise estimates are now provided in the NSSO documents for the Central sample only for important socio-economic parameters, the State sub-sample estimates are not available on time for a majority of the States. This results in getting less precise estimates, although the State sub-samples are being canvassed on a regular basis for a long period at enormous cost to the public exchequer. If the States are not in a position to provide State sub-samplewise estimates on time so as to promote pooling of Central and State sample estimates, effective steps should be taken urgently by the DoS to remove the bottlenecks. Alternatively, financial and manpower resources should be saved by curtailing drastically the canvassing of the State Sample schedules till the concerned States show their adequate capability for timely processing of the collected statistics.

The second major shortcoming is lack of longitudinal/panel data in respect of many complex socio-economic characteristics which would have provided more dependable estimates for planning and policy-making. The NSSO generates only cross-sectional data which are not adequate for studying the direction of change over time of several complex socio-economic characteristics.

like (i) gender-specific time use data on economic activities especially of female members of the households. (ii) income, expenditure and indebtedness, (iii) chronic unemployment and under-employment. (iv) households having members visibly deprived of "basic human needs", etc. The potentiality of Sample Registration Areas (SRA) (introduced by the office of the Registrar General of India under the Sample Registration System) in generating meaningful longitudinal/panel data as advocated in a recent article [9] needs to be considered carefully to strengthen the current national data system. For instance, longitudinal data generated from SRAs even if collected at regular intervals of 3-5 years would bring out more precisely the direction of change in the percentage of "visibly poor" (who are deprived of several basic human needs) over a period of years. A comprehensive analysis of estimates of poverty derived from cross-sectional data collected by the NSSO and longitudinal data that could be generated from SRAs would provide a better insight to the policy-makers into the real change in the incidence of poverty in different parts of the country.

**Computer Centre**

At present, the Computer Centre is catering primarily to the needs of the Department of Statistics. Rapid development of Information Technology has made the computer equipment currently in use by the Computer Centre rather obsolete and much less efficient as compared to those being used in many developed countries. Statistical analysis and dissemination of results consequently consume much longer time in India than what is feasible with more modern technology. The pre-eminent position in the field of statistics occupied by India in the 1950s seems to be gradually getting eroded mainly due to the widening gap in the information technology used in our country *vis-à-vis* other statistically developed countries. To start with, large-scale expansion of training and re-training programmes for the Indian Statistical Service (ISS) officers that would enable the new recruits as well as the present incumbents of the ISS to handle modern personal computers (PCs) with ease is considered essential for speedy dissemination and in-depth analysis of collected data. Re-training of staff at regular intervals is also extremely important for the ISS officers to cope with the latest advances in computer technology which are considered relevant for our country. For this purpose, assistance of the National Informatics Centre (NIC), Department of Electronics as well as other expert organisations including those in the private sector should be taken, whenever necessary.

**Department of Statistics**

A comprehensive review of deficiencies in our official statistical system is available in [1,7]. The remedial measures recommended in [1,2,7,8] should be carefully reviewed by a high-level Technical Committee to draw up a feasible action plan for strengthening the current statistical system.

Lack of proper career planning of ISS officers by the Department of Personnel in the initial stages and subsequently by the Department of Statistics has resulted not only in creating frustration amongst the majority of ISS officers but also in a deterioration in the quality of human resources engaged in statistical activities in the government offices. Emphasis should be laid on upgrading the capability of ISS officers through intensive training and re-training programmes and some other measures emphasized in [1]. The top post in the field of statistics in government departments should be filled only on a contract basis primarily on the basis of merit. Simultaneously, special efforts should be made by the DoS to upgrade the technical capability of those new recruits to the ISS who are selected mainly on "equity" considerations to raise their level of competence more or less at par with other direct recruits. For promotion to higher-level posts in the ISS, merit should get higher priority than equity considerations.

The National Advisory Board on Statistics (NABS) should be reactivated so as to play a much more effective *rôle* in advising the Department of Statistics from time to time to adopt a forward-looking approach for the development of official statistics in the country.

## Concluding Remarks

This short article has highlighted only a few areas of concern which need urgent attention of the DoS for revamping of the current statistical system. There are many other issues (outlined in the papers referred to above) which deserve attention from the DoS. With the revamping of the existing statistical system on the lines broadly indicated in the foregoing paragraphs, it should be possible for India to regain its pre-eminent position in the field of statistics in the years to come.

---

## References

1. Department of Statistics, Ministry of Planning (1980), *Report of the Committee to Review the National Statistical System*, Government of India, New Delhi.

2. K.C. Seal (1994). "Quality of official statistics and statisticians - the changing scenario after Mahalanobis", *Proceedings of Triennial International Symposium on Probability and Statistics*, University of Calcutta & Calcutta Statistical Association, Calcutta.

3. V.R. Rao (1993). "Statistics for rural development - an agenda for the nineties", *Journal of Indian Society of Agricultural Statistics*, April 1993, 1-47.

4. Arun Ghosh (1997). "India's crumbling data system", *Economic & Political Weekly*, June 21, 1439-1441.

5. K.C. Seal (1984). "National statistical system and quality of its output", *Presidential Address in Section of Statistics*, Seventy-First Indian Science Congress, Ranchi, January 1984.

6. Central Statistical Organisation, Department of Statistics, Ministry of Planning (1984), *Report of the High-Level Evaluation Committee to Review the Work of Central Statistical Organisation*, New Delhi.

7. K.C. Seal (1987). "On administrative restructuring for improved management of Indian official data system", *Calcutta Statistical Association Bulletin*, 36, 1-17.

8. K.C. Seal (1994). "National sample surveys for development planning - an overview", *Calcutta Statistical Association Bulletin*, 44, 239-250.

9. K.C. Seal (1998). "On utilization of sample registration areas for strengthening the national data system", *Proceedings of Twenty-First Conference of Indian Association for Study of Population*, Benaras Hindu University, Varanasi, February, 1998.

# INDIA'S STATISTICAL SYSTEM 50 YEARS AFTER INDEPENDENCE

## A. Vaidyanathan[*]

India can be justly proud of its system for collecting and disseminating statistical information on social and economic life of the country. At the time of Independence, a rudimentary apparatus was in place; but its scope and content were limited and patchy. The last five decades have witnessed a sea change : there are now sizable organisations, staffed by professionals, at the Centre and in the State governments, and in many of the ministries, to collect data on practically all important aspects of the economy and society. Considerable progress has been made in extending coverage, refining and standardising concepts, improving procedures, speeding up processing and publication. We have seen the widespread use of sample surveys and the creation of a specialised, autonomous National Sample Survey Organisation. As a result, a lot more information, of greater detail and coverage, is now available in the public domain than was the case at the time of Independence. Interaction between data gathering agencies and users has been fostered and in many cases institutionalised. Data and data-based analysis figure increasingly in debates on public policy. And this has stimulated critical discussion among professionals on the content, availability and interpretation of data. All this adds up to an impressive and commendable achievement. But one must also take stock, on this occasion, of several serious lacunae in the statistical system.

First, there has been considerable degeneration in the quality and timeliness of data. In agriculture, for example, land use and crop area statistics used to be compiled by village officials as the basis for collecting land revenue. As land revenue became less important and the tasks assigned to village officials became more diverse and demanding, the quality of land use and cropping data suffered. The situation has been aggravated by the abolition of traditional *patwaris/karnams*, and the weakening of periodic inspection of these records by higher-level officials. Crop yield estimates are based on sample crop-cutting experiments. While the sampling design is done scientifically, the responsibility for conducting the measurements in the sample plots is distributed between several departments. A supervision and sample verification mechanism is in place, but its effectiveness has become a matter of some concern.

In the case of organised industry, there is growing concern about coverage, reliability and timeliness of data relating to production of specific manufactured goods. A system of obtaining the data from industrial establishments had been built up. Though there was no legal compulsion, the dependence of industrial establishments on the government for various licenses and clearances induced them to furnish the data fairly regularly. With the dismantling of controls, and the DGTD, this inducement has disappeared. There is now a major problem of ensuring regular and reasonably complete reporting. The Annual Survey of Industries which is mandated by law, though comprehensive, takes a long time to be compiled and published thereby limiting its usefulness.

Instances could be multipled: large gaps remain in data relating to unorganised industry, education, health, trade, unorganised transport and services. Much of the available information is based on administrative reporting. Though sample surveys to get independent estimates of wider coverage have begun to be done by the NSS and the CSO, they are infrequent; reliability is also a problem. Inordinate delays in processing and publication of data is of course a general problem.

All this underlines the need to pay serious and urgent attention to revamping the system of primary data collection; exploring new techniques for both collecting and processing of data (such as more extensive use of sample surveys, computerisation of data processing and, in the case of

[*] Professor Emeritus, Madras Institute of Development Studies

agriculture, remote sensing), and ensuring that supervision and verification mechanisms do work independently and effectively. In some cases, legal sanctions will also be necessary; making it mandatory for industrial and commercial establishments above a certain size to provide specified data periodically.

The National Sample Survey, set up as an independent professional organisation to generate data on aspects not covered by government agencies, has made valuable contributions in improving data. It has demonstrated the effectiveness of sample survey techniques as a cost-effective and reliable way of getting information. In some areas - consumption, employment and assets - it is the only source of detailed and comparable data over time. The fact that its programmes, survey design and field procedures are reviewed and decided in close interaction with users and academics, is its strong point.

However as the organisation has become huge and depends on a permanent cadre of field staff, ensuring reliability and quality of field work has become more difficult. The efficacy of existing procedures for this purpose - supervision, inter-penetrating sub-samples, separate surveys by the Central and the State organisations using common designs, concepts and procedures - need to be reviewed. Reducing sample size through stratified selection procedures; simplification of schedules; research on techniques of field investigation to reduce the time and strain on informants - these are some of the areas which need attention. Of course, this will call for more resources. But resources will not be enough, unless there is a sustained effort, with constant interaction and feed-back from field staff and backed by continuing research, in a more intensive way than at present (in NSSO as well as academic and research/institutions) on survey design and methodology.

The tendency for user agencies in government to ask for more information, on more items and at a high level of disaggregation also needs to be checked. It is fashionable to ask for district or even block level estimates without being clear about the kind of specific questions sought to be answered and the information (in terms of content and precision) needed for the purpose. There is little concern about the effect of expanding the scope and the scale of these surveys on costs and on data quality. Unfortunately, the use of data-based analysis as an input to policy, though increasing, remains weak. This has serious consequences: the demand for data is indiscriminate; the relevance and utility of data are not subject to critical scrutiny by users, nor is there much feedback to the data-gathering agencies on these aspects; and there is little pressure on data-gathering agencies for improving quality and reliability.

Top managers of government agencies, in particular, should be encouraged to apply their minds more systematically to work out the data needed for various purposes, define their content more precisely and specify periodicity, disaggregation and level of precision with due regard to costs. The possibilities of utilising data already collected and available in the statistical system should be explored before taking up fresh surveys. There is an enormous scope for more detailed tabulations (including cross tabulations) and analysis of existing data. And in planning fresh surveys, a distinction needs to be made between "staple" data required to be generated at regular intervals on a comparable basis and those which serve special infrequent needs. While the former is properly the function of the NSSO, the latter can be commissioned as required by universities, research organisations and other agencies.

Users outside government have several complaints concerning lack of information on data sources and of getting access to the actual data. Thanks to the efforts of the Indian Econometric Society, the societies of agricultural and labour economics and other professional bodies, publications giving details of sources and limitations of data on various aspects are available to non-government researchers. The *Guide to Official Statistics* published by the government is also a valuable source. Some official agencies - notably the CSO - not only have non-official experts on their technical advisory committees but also publish details of data sources, concepts and

methodology underlying their estimates as well as the nature and rationale of changes made from time to time. (This unfortunately is an exception rather than the rule.)

The more serious problem is that even if users know what data is generated and where, getting hold of the data is a problem. Part of the reason is that all data are not published for public dissemination either on request or for a price. Inadequate publicity given to government statistical publications and the poor arrangements for their sale and distribution are serious problems. Efforts by organisations like the ICSSR, the FPW Research Foundation and the CMIE to collate and disseminate published data help to some extent. But a much larger and better-organised effort to expedite this process is essential. The idea of a data bank network mooted by the ICSSR is meant for this purpose.

Access to unpublished data, especially primary data, is more of a problem. Users complain about this. But it has to be recognised that the possibilities of mapping, analysis and interpretation based on published tabulations are not effectively exploited by researchers. In fact, they have not even scratched the surface of these possibilities. Nevertheless, it is true that certain kinds of analysis would need additional tabulations and/or access to primary data. Considerable bodies of primary data (e.g. individual census slips; schedules of various NSSO inquiries, corporate balance sheets, public finance and financial institutions), have been computerised and this practice is spreading. Until recently, government agencies were reluctant to take up fresh tabulations sought by researchers. Apart from a tradition of secretiveness in government, lack of computer capacity and high cost are also contributory factors. But the attitude is changing: the Freedom of Information Act is a sign of this. The Department of Statistics and the NSSO are also more willing for wider and freer sharing of unpublished tabulations and primary data. There is reason for optimism that the government will approve this policy.

But this is not sufficient to ensure wide and easy access to users all over the country. The creation of a wide network of regional data banks, where unpublished tabulations and unpublished primary data can be stored for ready access, is essential. Non-official organisations serving users must find the resources and be equipped to organise and manage data centres in a progressively competent manner. Moreover, the propensity to corner data is by no means limited to government: after all, control of information could be a source of immense power. This must be effectively checked in any non-governmental initiative. We need to create an environment in which the culture of free sharing of data will flourish.

# WHITHER STATISTICS?

## S. M. Vidwans[*]

I have been asked to express my views on the "direction in which the Department of Statistics (DoS) should work" and describe my "vision for the future of statistics in this country". A retired official should, however, be aware that his views might not be taken seriously. If he says now that something ought to be done, he could be silenced by the simple question as to why he did not do that when he was on the job. Similarly, an expression of his vision about the future would appear to be a trivial action of setting up objectives for others to pursue. Both are likely to be treated at best as dreamy abstractions and at worst as an advice without responsibility and commitment. I shall, therefore, mostly repeat here what I said, suggested, thought was possible and implemented, when I was a part of the system[1]. It is thus that I attempt to answer the question raised in the title of this article. I shall also point out why some of my ideas did not elicit a positive response. I could develop these views because I was not merely a collector of data but their user also for government's purpose, a rôle which, traditionally in my State, the Director of Economics and Statistics was required to perform. I must hasten to add that my observations are based on nine-year-old information; if the situation has since then improved, I stand corrected to that extent.

## 1. Review

It is tempting to look back over a span of 50 years and feel proud of the achievements of the Indian statistical system. It will, however, be incorrect to attribute the present level of development of the system to post-independence era. Much of the present official (administrative) statistics system in India was put in place before independence. Other landmark developments such as the decennial population census, or the application of sampling techniques for objective measurement of crop production or the continuous National Sample Survey were either achieved before independence or will soon be 50 years old. But the population census now produces less information, and with a greater time-lag, than it did in 1961; development of agricultural statistics meant merely a coverage by crop-cutting surveys of a large number of crops; and the NSS has yet to make much progress in timely release of its results. In the first three decades since independence, the statistical system developed mainly horizontally. Development mostly meant more statistics, more sample surveys, even more censuses, and a larger employment of manpower and resources in the system. The picture that emerged two decades ago was that of a system which had reached its "limits to growth" in its bulk[2]. Whether it progressed in its usefulness is, however, a moot point.

The period also witnessed centralization of statistical activity which must have had other merits but distanced the former collector and user of certain statistics from their production process. The State-level Censuses of Manufacturing Industries (CMIs) in the factory sector were taken over by the GOI as Sample Surveys of Manufacturing Industries (SSMI). The work of consumer price index numbers for, and the related family living surveys of, industrial labour in specified centres was taken over by the Labour Bureau. Fields, originally of State's interest thus passed on to the Centre and development of statistical activity in them depended upon the continued interest of the Central ministries. The same happened between GOI ministries. The Agricultural Labour Enquiry of the Ministry of Labour (MoL) was handed over to the NSS, but because the MoL does not show enough active interest in this survey of the most disadvantaged section of the society, it does not receive priority at the hands of the NSS. Centralization has not necessarily done justice to the cause of statistics in the related fields. We cannot perhaps reverse what has been done, but we should not centralize statistical activities any more.

[*] Formerly Director, Directorate of Economics and Statistics, Government of Maharashtra and former Member of the Governing Council, NSSO

The States too, I must confess, made a mistake. Like many other States' DESs, the Maharashtra DES (then called the Bureau) possessed an organization of its own for conducting sample surveys of its own interest, especially of important cottage and small-scale industries in the State. Subsequently, with the participation on a matching-sample basis in the NSS, which, with its multipurpose approach, held out a promise to generate data on all economic activities of the household sector, including manufacturing activities, the State's own survey organization languished. It was difficult to justify expenditure on its development when the DES had already created its NSS organization. The result was that neither could the NSS deliver the product the State's survey was producing nor, and more importantly, could the State's own sample survey capabilities be developed. Complete dependence on the NSS for the State's own information needs is a handicap the State now suffers from. Another fall-out of the shift of emphasis on sample surveys has been the neglect of administrative statistics, for which the government's administrative departments are equally responsible. (Examples: unemployment statistics and statistics of vital events.) Because sample surveys fail to substitute them, the net result has been a decline in the availability of quality information in several fields.

This brief review cannot end without a painful reminder of the commanding heights at which statistics was placed soon after independence. The CSO was made a part of the Cabinet Secretariat. It was a conscious policy decision that the highest statistical office should not function under the influence of any implementing ministry if statistics are to be produced objectively. Similarly, in the States, the DESs functioned under the finance or the general administration department. The situation has now changed, not exactly for the better. Over the last many years India did not have to look beyond its borders for statistical advice. In fact, a large percentage of statistical advisers of international agencies to other countries of the world were Indians. I understand that now things are reversed. The World Bank appears to be on the statistical scene of this country. I think all this does not augur well for the system. By routinizing statistical activity due to our rigid views, we have been partly responsible for this decline in the status of statistics. It is high time we recognize that the full potential of the usefulness of the gigantic statistical system remains to be exploited, make a serious review of what we do and change our orientation to ensure development of its usefulness. A possible line on which this can be done is suggested in the rest of this article.

## 2. Changes in Environment

**The Personal Computer (PC) Revolution** Computerization earlier meant the use of main-frame computers with their tremendous capability to process vast amounts of information with astronomical speed. Because of the huge costs involved, the economical way of processing the data required its centralization with the accompanying inefficiency of a centralized structure in deciding what to process, how to process and which processed product to disseminate. Any later changes in demand for information on the system entailed huge costs or delays and discouraged, in turn, such demands. The PC revolution should change all this. Since the processing of data can be decentralized, the tremendous advantage the PCs offer is the creation of capability in the system of processing the same data for information needs at sub-national levels, simultaneously fulfilling the national needs, and making the statistical system much more useful than hitherto.

**The End of Cost-Free Data Collection** We have taken for granted all these years, for the work of censuses, the ready, and almost free, availability of a vastly extensive field organization of lowest government functionaries reaching to the lowest unit of administration, deriving their authority from the government which guaranteed cent per cent response from the respondents. This created a certain casual attitude to taking up of census-type statistical projects; we have now the agricultural census, the livestock census, the economic census, the census of the poor. When, once in ten years, these government functionaries were asked to take up population census field work they took this additional work as seriously as their other duties. True, we use different functionaries: the *patwaris*, or the *gramsevaks* or the village primary teachers, for different censuses. But greater frequency

breeds contempt; with censuses now having been multiplied in number, and the government servants being an organized work force, the situation has changed. They are reluctant to do the census work, which has an adverse effect on the quality of the work they do, undermining the very purpose of taking a census. Unless absolutely necessary, we should not, in future, design statistical projects as censuses.

**The Reluctant Respondents** We have similarly taken for granted the favourable "response" factor in our census and survey operations. The respondents of the questionnaire, aware of the silent authority of government that backs its surveys, had been generally co-operative in providing information. The psychology of a "subject population" endured for some decades after independence. Impervious to the limits of their patience and willingness to co-operate, we use unreasonably long schedules in our surveys. People have changed and become aware that they are not bound to provide information to government surveys. Cases of non-response or bad response will increase. We should seriously review our survey-taking techniques. I had experimented with simple methods of data-collection and the results were encouraging. An interested reader may refer to my two papers[3],[4]. Statisticians seem to think that for collecting data on any subject, all that is necessary is to design a household schedule and use a household sample approach. It has failed in the surveys of trade and services, but we have not developed any other approach to handle these sectors. We seem to be so much wedded to the household approach that we have lost ingenuity to find solutions to our problems.

**End of Planning Era?** Statisticians have been conditioned for long to believe that the only purpose of statistics is to be useful for the needs of planning. I had alerted one of the DGs of CSO on several occasions to the danger of making development planning the sole *raison d'être* of statistics, and to the implicit neglect of the larger purposes of statistics. For, if the government decided to curtail its rôle in development, the gigantic edifice of the statistical system would lose its only props. This was not a doomsday forecast. The recent change in policy, about the rôle of government in the economic management of the country, including planning, has proved me right.

## 3. Statistical Information System

**The User of Statistics** Government statisticians should at least now wake up to the basic purposes of their statistics and their system. In many a modern democratic country, it is the people's right to receive and the government's duty to provide objective information about the conditions of their lives, as the government alone can afford to do so. The implications of this obvious requirement seem to have been lost in the blinked view held so far of the purposes and uses of statistics. The first implication is the variety of facets of life - economic, social, political - on which information must be produced. The statistical system does not suffer from any serious deficiency on this count as far as its coverage is concerned. Improvement in its quality should continue to be its all-time concern. The second stems from the fact that the institution of government works on many levels: the national level, the State level and the level of local self governments of municipal corporations, municipalities, *zilla parishads* and *panchayats*. The statistical system must produce information at all these levels. The third implication is that whatever aspects of people's life it brings within its fold of "management", the government (at all levels) needs quantitative information for good management. And the same information must be available to the people to enable them to judge the government's decisions independently. All this underlines the necessity of a decentralized system of information generation. It also clearly defines the government, at all its levels and in all its branches, as the prime user of the system. The restrictive definition does not negate the class of researchers and academics in the many institutions of research and learning in the country, as users of statistics. Their demands can be met by a separate arrangement. But it should not prove a hurdle, as it sometimes does by pre-empting resources, in meeting the needs of the prime user, the government.

**Types of Information Needed** One may distinguish information needs of the government according to the two types of target of its decisions or policies. The first comprises the sectors of

economic activity; here much of the needed information is currently available. The other comprises certain specified groups of people. For both, information should be available by the location: location of economic activity, and for groups of people identified by their residence: rural/urban, administrative or geographical, or agro-climatic regions. In the case of people, what is also important is their identification as groups by socio-economic characteristics such as landless households, agricultural labour households, cultivators growing certain crops, or cultivators having irrigation facilities, housing conditions of households, social class of households, etc. These characteristics are such that they clearly identify households or people for government's policy decisions into operationally meaningful classes, which quality the synthetic classes based on monthly per capita expenditure (MPCE) or the fractile groups do not possess. To enhance its usefulness, the statistical system should therefore produce information about economic activities and people for different regions at the sub-national level, and about people by their socio-economic (SE) groups. I shall call both these statistics for micro-level use. The following discussion suggests how this needs to be done.

**Change in Orientation of the Present System**  The qualifying adjective "statistical" of the term "information system" is important. For information system is a wider concept, not necessarily requiring exact measurement and permitting subjective interpretation of information. That is not what the statistical system should be converted into. The concerns of the statistical system, viz. collection, derivation and presentation of objective statistics that pass the test of validity, based on sound statistical principles and standards, must remain paramount and should not be forsaken. But, its orientation should now change for better decision-making by governments and their organs, and that will require a change from the "fixed-product" approach that is inherent in the process to "demand-determined-production" approach, which advances in the data-processing technology have now made possible.

The Indian Statistical System is partly decentralized and partly centralized, with the States being responsible for the first part. The national-level data for sectors which are "State Subjects" are obtained by a process of aggregation of State-level data. The generation of information at the lower levels of administration is not a problem and is the responsibility of State governments. One area deserves special mention. Given the fast increasing urban population, the highly urbanized States should treat urban statistics as a priority sector for statistical development and should now vigorously implement the scheme of publishing Municipal Year Books, started many years ago, which has suffered a neglect.

It is the centralized part of the system which is the focus of attention in this article. Two levels of use of these statistics should be clearly recognized: the macro-level and the micro-level. Most of the use of these statistics for estimating national-level parameters and analysis falls in the first category. In my opinion, it is the micro-level where the great potential for making the statistical system useful lies. Provision of information at the micro-levels is in this case the responsibility of the Central government. But, as I point out below, the concern of that government with only the macro-level use of data has led to the neglect of the needs at the micro-level. That is why a change in the orientation is required.

**Population Census**  The Government of India (GOI) on the one hand encourages decentralized planning from the district-level going down now to the level of *panchayats*. In particular, it used to ask the States to prepare district-level manpower plans! The CSO also exhorts the States to improve their estimates of State income and work out district income. But the single project of creating economic information at the lowest administrative level, implemented once in ten years, has stopped giving the information that it has collected at that level. Since 1971, the Census has failed to produce data on the count (number) of non-agricultural workers by detailed occupational classification even at the district level. What it produces are "estimates" which are not usable at that level. The irony is that this has been a consequence of the use of improved technology of data processing: first the decision to use computers, and then, because we did not have enough resources

for data-entry, the use of sampling of collected information! The result? The modernization of the process cannot now produce information that used to be generated by hand-sorting till 1961. And think of the waste of the colossal amount of information already collected! Now, with the cheap availability of PCs, is it inconceivable for the population census to do what it ought to do and it alone can do, namely, to give actual numbers (not estimates) which provide crucial economic information at the lowest administrative level?

**Agricultural and Livestock Censuses**  The agricultural census fails to produce data with cultivator as the unit of information, and so does the livestock census with household as the unit. Little thought is given to the immense amount of resources - perhaps because some of them come free - spent on the two censuses, the end products of which are aggregated figures of area by different classes and numbers of livestock of different kinds. Worse, it is not realized that the separation of the subjects of land and livestock holdings is a totally unrealistic way of dealing with the most important sector of the economy where cultivation of land and animal husbandry are inseparable parts of an integrated single economic activity. The irony, again, is that the same GOI Ministry of Agriculture carries out the two independent censuses and has been knowingly blind to the simple task of integrating the two censuses. In fact, it took several years of sustained pursuit of the idea on my part to bring home this simple truth. That resulted in the appointment of a Committee to implement the suggestion. But I do not know whether the two censuses have been integrated or continue to remain separate.

The agricultural census also illustrates a serious case of missing something simple but vitally useful for the government in the mechanical thinking about a project. The only indicator available of inherent productivity of land is the rate of basic land revenue (at least in a large number of States). Although the census is no more than a census of land records, it did not include this simple item of recorded information. Considerable persuasion on my part finally made the Maharashtra census include it in 1985. The Maharashtra census also showed how the same census operation can be used to collect important items of information without extra cost. It collected data in 1985, from the records, on number of trees in the operational holding for "fuel" and "horticultural" purposes separately, and in the 1990 census, on classification of holders as males and females. Most importantly, it carried out my suggestion to tabulate the census results on a sample basis, to generate information for different classes of cultivators. I quote this merely to illustrate how the utility of a statistical project can be enhanced if its designers are sensitive to data needs of the government.

**Official Statistics of GOI Ministries**  As said above, the CSO asks the States to improve their statistical systems, and especially the statistics of State income. On the other hand, centrally collected information required by States for these purposes are not made available to them even though the data have been compiled by aggregation of sub-national data. Two examples will illustrate the point. The data on transport of goods and passengers from each of the railway stations within the States, though available with the Ministry of Railways, are not given to the States' DESs in spite of computerization of the railway statistics. Similarly about income tax data. It was easy earlier for the States' DESs to obtain from local income tax offices data on income tax collected from their regions and cities. Later, apparently a policy decision directed income tax offices to stop giving these data to the States. But this unfairness did not prevent the income tax department from asking the States to provide it with lists of motor vehicle owners! The two cases are symptomatic of two institutional traits: (1) the sense of ownership of data prompting the exercise of any arbitrary power about data dissemination to the extent of preventing other State organs from using them; and (2) the lack of trust of Central ministries in the State governments' use of the data. This carries an offensive smell of politics to the field of statistics. In a field which appears in the concurrent list of subjects in the Constitution, the sharing of data by Central ministries with States should be as mandatory as it is for the States to furnish data to the Centre.

**National Sample Survey**  In the case of the NSS, the problem has a different character. A number of States participate in the NSS on a matching-sample basis. Unlike in the above illustrations, the

States do not have to depend upon the NSSO for their data needs, for they can use data from their own matching sample. But the matching sample arrangement dictates that the States should produce tables according to the NSSO tabulation programme (TP), and that massive TP takes up most of their resources preventing them from producing information on demand, unless they make special efforts[5].

The NSS should review the very concept of a catch-all TP and the associated idea that once its tables are produced, the use of NSS data and the responsibility of the NSSO are over. A vicious circle seems to operate here. The fear that unless a table considered useful (by one or more "users" on the GC) is put in the TP, it might not ever get prepared on demand, tends to pack as much as is conceivable in the TP bloating its size to thousands of tables. The mammoth size of the TP, in turn, ties down processing resources for this single objective, making it impossible for the NSS to produce any "on-demand" tabulations, and thus reinforcing the tendency of users to "put it in the TP if you will ever want it".

There is no denying that one should have a basic TP. The question is about its orientation. In line with what I perceive to be the information needs of the government, the basic TP should contain Statewise - and where sampling errors permit, regionwise - tables showing the estimates of the subject parameters, and where relevant, their distributions according to SE classes of households. The household classification based on monthly per capita expenditure (MPCE) or fractiles is relevant for macro-level analysis and policies but not for policies of which different groups of people are the foci. People do not carry MPCE or fractile identification as they carry identification by their SE classes. Further, I do not see much point in the same type of macro-level analysis by MPCE or fractile classes for each round. Analysis by SE classes will bring out changes in the conditions of different SE groups over the years. Further, if the basic TP contains only such simple tabulations, its bulk will be reduced considerably. I had once demonstrated, by actual tabulation of consumer expenditure data for Maharashtra, how these simple tables - which go a long way in being useful - can be speedily produced, cutting down the gigantic size of the TP by a factor of 10!

**Aversion to Change**  The suggestion does not find favour with the NSSO-GC. I think the reason behind it is the difference in perception of what a user like me, appreciative of the needs of the government, considers useful, and of those more interested in macro-level analysis of one type or another. The problem about the NSSO is that it is not a user of its product, is accustomed only to the needs of the macro-level analysts in the GC, and does not appreciate the significance of any other use of the NSS data. Since what the World Bank thinks seems to have assumed importance for our statistical system now, I wonder whether we are waiting for their recommendation of this shift of focus of the TP on SE classes. What I am asking for is something they have already done in their Living Standards Measurement Studies (LSMS). The usefulness of the NSS is beyond question. But that lets an important point left unsaid, namely, that, in relation to its vast potential of becoming useful, it remains considerably underutilized because of a rigidity of view about what it can produce. I hope that at least the States' DESs, which participate in the NSS on a matching-sample basis, would see the immense opportunity they have of increasing the usefulness of their NSS data for the purposes of their governments.

The same types of reason explain why much of what I have suggested above should be done does not get done. The Central ministries have a limited view of usefulness of data and are unfamiliar with the States' needs. Unfortunately, the statisticians in the corresponding State departments also have a routine rigid view of use of their data. This stems from their relationship with the Central ministries, a factor I shall examine shortly.

## 4. Related Issues

**Data Synthesis, Linkage and Bank** The above discussion has indicated how the utility of each statistical project can be enhanced. Quite frequently, all the required information is not available from the same survey or census or other source, but is spread over a number of them and needs to be synthesized. This is an area in which research needs to be done to develop appropriate techniques of statistical analysis of such dispersed data to create the necessary information. Secondly, we tend to look at individual statistical projects, be it the NSS or population census or any other census, in a disjointed manner. If among the data elements of a survey or census, certain key items of information of other surveys are included, it will establish a linkage between the fields of data covered by different projects, increasing the totality of information content of all projects much beyond the sum-total of their individual contents. For example, on my suggestion, the NSS consumer expenditure schedule includes information on certain items (about households), which fall in the domain of other surveys, such as cultivation of major crops or ownership of livestock. Given the creation of computer capability for analysis, one can visualize the great potential that will be created by interlinking information from the diverse sources in this manner. Lastly, as a starter, it is essential to create a system to store all important data in a data bank - even unit level data on sample basis, where necessary - and keep it accessible to the users. Computerization of the statistical system should set this also as one among its important goals.

**Coordination** An important function of the CSO is the coordination of statistical activity in the country. This has largely meant coordination between the CSO and the States' DESs. But it is the coordination between the statistical departments of Central ministries and the CSO that needs to be improved. The ministries seem to be deciding things for States independently of the CSO or the States' DESs. Two cases have been pointed out above. Another example was the sudden decision some years ago of the Education Ministry to stop one of the two forms which schools had to file with the State's Education Department. And incomprehensibly, it was the form which was to be filled in September, just after the school begins, that was dropped, defeating the very purpose of obtaining advance information about the educational situation, which is as much valuable as that obtained at the end of the school year. In the present system, the work content of the statistical activity in the State departments is decided by the corresponding Central ministries. There is no lateral linkage at the Centre nor at the State. The implied accountability only to their Central ministries has endowed the departmental State statisticians with only a narrow perspective of their work and little capability to appreciate the wider significance of their data. This is certainly not a state of affairs conducive to the improvement of usefulness of the statistical system as a whole.

**A Caution and a Hope** With the ready availability of software packages for statistical analysis, the analytical statistician has almost become redundant. The new concept of information systems coupled with computer technology could do the same to government statisticians. The orientation of the government statistical system must change now towards producing information for government decision-making. What it means was explained by me in my paper in the *Journal of IARNIW*[6] also. Unless this happens, government statisticians will be reduced to mere collectors of data and the work of using data for government's management purposes will be taken over by the new breed of information experts waiting in the wings.

This should not happen, and hence this article. The essence of what I have suggested needs a system of decentralized processing of data to transform them into information at sub-national and micro levels. It would imply that the present centipede-like structure of the system, where a tiny part does the brainy technical work and the multitude of statistical functionaries do absolutely routine work, must change. We need now a decentralized organizational structure with the regional or local units acting like microcosms of the parental organization. The change will bring in its trail considerable job satisfaction to the statistically qualified men and women working in the system, giving them the important rôle of users of information for the micro-level analysis at the local level.

To the young qualified statisticians, it would offer a chance of using their acquired talents and make for an intellectually exciting career. The total qualitative content of the system will improve and only then would the system remain continually alert to the changes in environment that take place and adapt to them.

---

[1] I presented some of these views in the biennial conferences of State and Central statisticians in the 1970s and 1980s. I set down these ideas concretely first in my paper "Information system for policy decision-making at the State level", submitted to the Nainital Seminar of the Indian Association for Research in National Income and Wealth (IARNIW) on Monitoring the Current Trends of the Indian Economy: Data Base, in 1977, a paper which was published in the *Journal of IARNIW, Vol. 2*, No. 2 (April 1978). I submitted these views again as a member of the official Working Group appointed by the GOI on the statistical schemes for the Sixth Five Year Plan. The third occasion arose when I submitted a paper titled "Official statistical system of States in India: a line of the future development" in 1979 to the Statistical Review Committee appointed by the GOI. The last occasion was a paper I submitted in 1985 to the NSS Seminar at Calcutta on the Impact of Computerization on the National Sample Survey.

[2] For, it was almost difficult to formulate any "new" statistical schemes for the Sixth Five Year Plan, except, of course, for computerization.

[3] "A place for quantal information in socio-economic surveys", *Sarvekshana*, the Journal of NSSO, Government of India, April 1978.

[4] "Use of quantal method for estimation of agricultural production", *Recent Advances in Agricultural Statistics and Research*, (ed. Prem Narain *et al*), 1991.

[5] The tabulation work of the first Debt and Investment Survey of the Reserve Bank of India, carried out through the NSS, was centralized in the RBI. It so happened that the simple and direct table on the distribution of outstanding loan according to its source was not included in the first priority tables. Two years after the survey, when the State level Committee on the liquidation of debts of the poor wanted this tabulation for examining the importance of private moneylenders, the Committee had to be told that the table would be available one year later! We got the data in the particular blocks of the State sample schedules punched and mechanically tabulated-- this was before computerization-- to meet the simple needs of the Committee. The example illustrates that the representation of users on the NSSO-GC is no guarantee of capability of the NSS to produce useful results. It also shows how a pre-determined TP becomes a hindrance to producing results "on demand".

[6] See (1) above.

# INDIAN STATISTICAL SYSTEM FOR THE 21ST CENTURY

## Pravin Visaria[*]

A robust statistical system is both the foundation and an important correlate of development of a country. It is not just an outcome of the penchant for quantification, or the need to describe before we can prescribe. It reflects also the need to evaluate the ongoing processes of change and to identify the shortcomings or problems that must be corrected speedily in any civilised society. Besides, the figures that form statistics relate to the live figures, whose welfare and progress are the primary objectives of a democratic society and polity.

The current year, 1997-98, is being celebrated to mark the golden jubilee of India's independence from British rule. Two years from now, in 2000 A.D., the National Sample Survey Organisation, which has been an important pillar of the Indian statistical system, will complete 50 years of its existence; and it will be time to celebrate its golden jubilee. Such celebrations must also be the time for introspection to review our achievements and progress and to identify the deficiencies to ensure that we build on the former and overcome the problems and can compete with the best in the world.

### Stylised Facts

As I reflect on these issues, the following stylised facts strike me as an apt description of the Indian situation. First, India was once considered a statistically developed country; but over the past few decades, we have probably lost our privileged status. Secondly, the database of the Indian economy has probably deteriorated, particularly in regard to agricultural and industrial statistics but to some extent even on other themes. Thirdly, the country has gradually begun to catch up with the rest of the world with respect to information technology, but the progress is painfully slow, certainly slower than can be expected of a country with the indigenous scientific capacity to develop nuclear weapons. Fourth, there has been a welcome change in the recent past in the policies about giving researchers an access to the primary data collected from the households or establishments; but the efforts relating to the dissemination of the results of the surveys are woefully inadequate.

Each of these statements needs some elaboration. In effect, we need to note some general aspects of the Indian setting, that limit our ability to achieve full-scale development of the statistical system. It is essential to separate the extent to which the leaders of the statistical system can be held responsible for the weaknesses of the situation from the systemic constraints. Without a judicious effort in this direction, we could take actions that would prove counter-productive and further damage the already weak system.

### Statistics and Statisticians

Sometimes one begins to doubt whether India ever was a statistically developed country. The Statistical Adviser to the Prime Minister of India, Professor P. C. Mahalanobis, did indeed preside over the United Nations Statistical Commission in the 1950s for several years. However, he had gained his stature by his constant efforts to solve the problems of inadequacies of the Indian Statistical System through sample surveys and his essentially empirical or experimental approach to

[*] Director, Institute of Economic Growth, Delhi and Chairman, Governing Council, NSSO

the resolution of problems. Also, with his wide interests, the presence of scholars from different disciplines at the Indian Statistical Institute enabled Professor Mahalanobis to interact with experts with varied backgrounds. The location of the National Sample Survey in the Indian Statistical Institute probably fostered an academic environment in which methodological studies and experiments could thrive. Several Indian experts served as advisers to help develop the statistical systems of other developing countries, but whether and how far they continued academically creative work thereafter has not been assessed.

Yet, given the widespread impression of a deterioration in the situation, our statisticians need to reflect individually and collectively on the factors that have contributed to the erosion of the standing of Indian statisticians over the years. Quite possibly, the problems are not unique or limited to the statisticians. The decline in the academic standards that has accompanied the expansion of universities and research institutions is ubiquitous and is evident both in the social and natural sciences. The brain drain, the exodus of the most promising students to the universities abroad for higher studies, and the progressive decline in the inter-state mobility of scientific and technical personnel have also had an effect on the situation. The efforts in the early 1970s to establish some "centres of excellence" in every State have also not succeeded to the expected extent; some even say that these efforts have failed. Of course, once an institution is established in India, the almost inevitable vested interests prevent its closure, even if it fails to achieve its objectives.

Turning to statistics, there is an unfortunate unwillingness to pay the real costs of collecting good reliable statistics. The difficulties of field-work in rural as well as urban India are not adequately appreciated; and as a result, the quality of data inevitably suffers a great deal. The detailed schedules prepared for our surveys of consumer expenditure or for estimating the value added by the household or small-scale enterprises are conceptually sound; but they overlook the ground realities and limitations of the respondents, very few of whom maintain accounts of their expenditure or income flows, both at home and in their family enterprises.

Given the size of the country in terms of both area and population, with its spread over nearly 600,000 villages and 3,700 urban agglomerations, the problems of data collection are quite serious indeed. According to the 1991 census, nearly 67 percent of the 581,000 villages in India (excluding Jammu and Kashmir) had a population of less than 1,000 persons each. Some 51,000 villages had a population of less than 100 persons each. These small villages tend to have a dispersed population. Not many of them get selected for our national surveys, because the selection is based on the criterion of probability proportional to size in terms of population or the number of enterprises. The lack of facilities of accommodation and food, etc., limits the ability of investigators to devote themselves to the assigned tasks in most rural areas.

There has been widespread concern in recent years about the evident underestimation of our national income originating in different sectors, particularly the small-scale industry or the services sector. In this respect, I am reminded of a statement of M. W. Yeats, the Census Commissioner for the 1941 Census. What he said about the census applies also to the surveys and accordingly, I shall modify it. In his Administration Report, Yeats wrote: "no census (or survey) ... can rise above the level of the people... . In any census (or survey), the citizen's part is the more important and fundamental, for our enumerators (or investigators) are primarily scribes and it is the citizen's collaboration, which produces the valid record". As is well-known, many highly successful self-employed workers prefer to understate their earnings or surplus. The survey investigators are unlikely to receive the requisite co-operation from these well-off people; and it would be a serious mistake to assume that they are less clever than the survey designers. This systemic constraint has to be recognised, because unaccounted incomes are reported to be a fact of life in the developed

countries as well. Of course, there is no valid excuse for using the census data on workers, which are known to be gross underestimates, for estimating the value added in the unorganised trade or the other services sector.

## The Deterioration of the Database of the Indian Economy

It is necessary to assess carefully the widespread impression that the database of the Indian economy has deteriorated in recent years. A frequently-cited example relates to the registration of births and deaths, which is said to have become less complete than it was prior to independence in many of the former provinces of British India. A knowledgeable colleague has attributed the problem to the casual attitude of the senior officers such as the collectors, who earlier used to check the records of births and deaths whenever they visited a village. It is also argued that after the development of the Sample Registration System, the civil registration system has been neglected. One indication of the problem is the slackness in the timely publication of whatever data are collected. If these data can become reliable, the inter-district differences in the level of fertility and mortality could be identified and could permit proper policy responses and programmes.

The other examples of possible deterioration of the database relate to the estimates of area under different crops and the production of different industrial commodities. (These data determine our estimates of agricultural output as well as industrial production, and therefore, their contribution to the national income.) The former has been affected by the reported increase in the work-load of the village level revenue officials in the wake of the introduction of many rural development programmes. The alternative estimates could be built through remote sensing satellites, but the promise and problems of this alternative approach need to be assessed and confirmed through field studies. There is an urgent need for close collaboration between the interested experts of the Indian Space Research Organisation and the authorities engaged in the collection of data relating to agricultural production.

The data on industrial production have deteriorated because with the liberalisation of the economy and the removal of many controls and licensing requirements, the industrial firms have lost some of the incentive to report their production to the government, i.e., the Ministry of Industry. The problems are compounded by the difficulties of obtaining records relating to the establishment of new industrial units, although some co-ordination with the power supply authorities could permit preparation of a good frame of all the industrial units using electricity. Of course, the listing of industrial units not using power remains a problem.

## Potential of Information Technology

One favourable development that can facilitate remedial action to improve the database is the considerable progress in the field of computers. The costs of data processing equipment have declined sharply over the past five years and powerful computers can now be purchased at a fraction of the earlier prices. The data entry can also be accelerated; but in addition, there is the possibility of eliminating the need for data entry by using palm-top computers for the collection of information from the basic respondent units. Unfortunately, according to the available information, the experiments in this regard have not been conducted with the requisite care, with due attention to the questions of validation and the comparability of data collected through alternative approaches. Some countries of the world have begun to rely on more powerful laptop computers for data collection in surveys; and we need to rethink whether, in view of the steady decline in the costs of even the laptop computers, a cost-effective alternative to palm-top computers is really beyond the realm of

feasibility. Of course, due care will have to be taken to ensure that non-availability of electricity or the batteries does not frustrate the planned schedule of work.

Not many users of data are aware that there has been considerable improvement in the speed of processing of data gathered in our surveys or censuses. This unfortunate fact was brought home to me at a recent meeting of the National Advisory Board on Statistics by a remark of a former Secretary of the Department of Statistics about the long time gap between the collection, tabulation and publication of data. Obviously, our dissemination of the information about our activities is not satisfactory.

The 1991 census has taken a significant step ahead in making the primary census abstracts for every village of India available on floppy disks, within about three years of the completion of the count. The data make varied interesting analysis, such as was impossible in the past. This experience has promoted interest in putting the data of the PCAs of earlier censuses also on the electronic media. Many users of the census data are concerned that if they do not have access to the personal computers, they remain at a disadvantage, because the data are not published in book form. This is a valid complaint, and there is really no excuse for some of the decisions of the census authorities, which have delayed the release of the 1991 census data even on floppy disks to the end of 1996 or early 1997.

## Importance of Timely Dissemination of Data

As noted above, the advances in information technology have made possible relatively speedy release of the statistical data. However, we need to review our procedures of data processing and examine whether it is really necessary to write new computer programmes every time we want to generate a fresh table. Other countries are able to use "canned" programmes with advantage, and we cannot be an exception. Likewise, despite acceleration of the speed of release of the national data and tables, the NSSO has in recent years given up the earlier practice of releasing State-level data in separate volumes. This is a retrograde step, and the State-level tables must be released through electronic media. Equally essential is the quick dissemination of the final reports. The NSSO fails to derive well-deserved recognition for its labours, because it does not spend enough time or resources on the press briefings and distribution of summary reports. The suggestions on these matters have not elicited adequate response or action, despite the fact that not much expenditure would have to be incurred for the purpose.

## Vision for the 21st Century

To improve the present situation, we need to increase the interaction between the producers of data and the analysts in the academic world. We also need to give the members of the Indian Statistical Service (ISS) adequate opportunities to remain up-to-date not only in regard to their own discipline but also about the applications of data in real-life situations. For this purpose, it is essential to revive and forge close links between the statisticians in the government and the academics in the Indian Statistical Institute and other scholars interested in the allied disciplines.

It is important to recognise also that the statistical system does not gain strength in hierarchical systems of command and discipline. Free and frank discussion about all problems is absolutely essential for progress of a statistical system. As noted above, no individual group can be blamed for the systemic constraints. Most simplistic solutions are unrealistic and one cannot earn respect by hasty compromises on matters that merit detailed methodological work. The Indian Statistical System certainly has the potential to retrieve its lost glory if we continue frank dialogue and initiate necessary actions with hard work and dedication to the goal.

# SOME PERSPECTIVES ON STATISTICS IN INDIA

## I. V. Basawa[*]

## 1. Introduction

This is a personal, and perhaps a naive, overview of some current trends and future directions of Statistical Science in the Indian context. India has a long and rich tradition of original research contributions to statistics as a discipline and its applications in a variety of fields including government (official statistics including sample surveys), agriculture, industry, epidemiology and ecology. The Indian Statistical Institute pioneered by Professor P. C. Mahalanobis deserves the major share of credit for the past and current success of Statistical Science in India and abroad. The Department of Statistics of the Government of India, in the last fifty years, has made significant contributions to the organization and development of statistical applications in various areas affecting the nation. The Central Statistical Organization (CSO), the National Sample Survey Organization (NSSO) and the Computer Centre (CC) are the three pillars of the Department of Statistics. Each of these three organizations must be congratulated on its success in achieving the goals which benefit the country, in spite of the scarce resources. I am particularly impressed by the achievements of the NSSO and the CSO. The Computer Centre (CC) assists the NSSO and the CSO in processing their data.

## 2. Current Trends

Statistical Science is basically concerned with collection of data, analysis and interpretation of data pertaining to a given project. Both the NSSO and the CSO have excellent track records in the collection and analysis of data on major studies that affect the nation. The question that arises is: How are these studies interpreted? Other related questions are: do these studies lead to policy decisions or recommendations? If so, how successful are these recommendations at the implementation level by the Government? Is there a follow-up mechanism at the Department of Statistics which oversees the implementation stage and makes cost-benefit studies?

After collecting, analyzing, summarizing and recording of various official statistics at enormous cost and efforts, it would be a great waste if the findings of these studies are lost in the bureaucracy and are dumped into official archives which are largely ignored by the government and the politicians. Even though decision-making may not be strictly within the jurisdiction of the Statistics Department, it behoves us to think seriously about the enormous wastage of official statistics if the decision-making aspect is completely ignored.

Some describe statistics as a science of making decisions under uncertainty. As an external observer, without the benefit of the experience of a government statistician, it is my opinion that most government statisticians are denied the rôle of a decision-maker. I hope I am wrong on this belief. If I am right, the question that arises is: how does one improve the situation?

Recent technological developments have drastically changed many scientific investigations. Statistical science is certainly not immune to these changes. As a matter of fact, statistical methodology has been undergoing major changes owing to the recent computer revolution. It is unfortunate that India has been lagging behind in the utilization of computer technology in statistics.

Statistics is an interdisciplinary science. It is very important for statisticians to interact with various sciences, e.g. social and political sciences, agriculture, ecology, education, food science,

[*] Professor, Department of Statistics, University of Georgia, Athens, GA, USA

engineering and industry. Has the Government of India Statistics Department explored the possibility of joint projects with other departments in the related areas mentioned above? I am glad to see that both the NSSO and the CSO have been working on projects in agriculture, human resource development, housing, economic indicators, environment statistics, etc. If some of these projects are done jointly with various departments of the government primarily responsible for these areas, and the findings are properly co-ordinated, one can make progress in developing fruitful interactions with other disciplines. This effort would also educate the statistically illiterate people in other departments on the benefits of statistical thinking in their individual departments.

## 3. Future Directions

Here are some of my naive thoughts and suggestions for future directions for Statistical Science in India.

### (a) Balance between Theory and Applications

Professor Mahalanobis always emphasized applied statistics and stressed that any statistical theory should have its roots in a real life practical problem. Dr. P. V. Sukhatme was a pioneer in applied statistics and sampling. In later times, unfortunately, the emphasis shifted to more theoretical developments both in India and abroad (especially, in the United States). At some stage, the situation went so far that most statistical research appeared purely as a mathematical artifact. More recently, however, there appears to be a general consensus that we need a balance between theory and applications. This can be achieved by interaction between statisticians and scientists from various other disciplines.

### (b) Utilization of Modern Computer Technology

The use of statistical computation in the analysis of data is almost universal in the developed countries. Because of some short-sighted policies of the past and pressure from some political groups, Indian statistics fell far behind the developed countries in catching up with and benefiting from computer technology. The unfortunate mistake should be corrected. I am glad that more recently, Indian statisticians in the government, private sector and the universities are realizing the benefit of computers. The Computer Centre of the Department of Statistics is a welcome addition. Its rôle , however, should be greatly enhanced and it should have branches throughout the country.

### (c) Collaboration between Government, Private Sector and Universities

In order to keep up with the current developments in statistical methodology, both the government departments and private industry can benefit from interaction with graduates students and faculty from the universities including the Indian Statistical Institute. The universities on the other hand can also benefit from such an interaction. In the United States, such interaction is encouraged by the relevant government departments by providing summer fellowships and internships to the university faculty and students. Many industries in the US offer summer internships to the students and consult the University faculty on their projects. Various government funding agencies support basic research activities in the universities. Indian Statistical Science can benefit by following some of these initiatives.

### (d) Training and Education

India has played a pioneering rôle in respect of training and education in statistics, thanks to people like P.C. Mahalanobis, C. R. Rao and the others. At one stage in history, Indian statistics was at the forefront of statistical research in the world and, indeed, played a leadership rôle . Unfortunately, this cannot be said about the current situation regarding the *rôle* and status of Indian

statistics, even though I hasten to add that there are many good and internationally well recognized statisticians in India today. The deterioration of the international status of Indian statistical research has been due to many reasons including the well-known "brain-drain". While India does have a good infra-structure for training and education, there should be renewed efforts to improve the quality of education.

## 4. Concluding Remarks

Interdisciplinary studies, decision making, collaborative studies between government, industry and the universities, balance between theory and applications, and improvements in the quality of statistical training and education are some of the suggestions offered for the future. More general recommended references in this context are Cox (1997), Pullinger (1997) and Rao (1989). I take this opportunity to congratulate the Department of Statistics for its various achievements over the last fifty years.

---

## References

Cox, D. R. (1997). "The current position of Statistics: a personal view", *International Statistical Review, 65,* 261-290

Pullinger, J.(1997). "The creation of the office for national Statistics", *International Statistical Review 65,* 291-308.

Rao, C. R. (1989). *Statistics and Truth.* Ramanujan Memorial Lectures, International Co-op. Publishing House, Maryland.

# MY ASSOCIATION WITH THE NSS - REMINISCENCES AND OBSERVATIONS

## Nikhilesh Bhattacharya[*]

### Early Recollections

I joined the Indian Statistical Institute (ISI), Calcutta, in a part-time capacity in May 1952, and later, in a full-time capacity, in December 1954. The NSS had been started in 1950 and the NSS Department of ISI served as its technical wing from 1950 to 1972. I never belonged to that Department. I worked in a Unit called the Special Technical Unit, directly under Professor Mahalanobis, from 1955 to 1959. But I was drawn towards the NSS, thanks to the generosity of the stalwarts of the NSS Department of ISI, who invited me to become associated with the high adventure which the NSS was at that time.

First, they got me trained in the ISI methodology of crop-surveys and sent me to different State headquarters to impart training to the NSS field workers just before the NSS 13th round (1957-58) crop estimation surveys were started. Later, I was a member of the ISI team when crop-cutting experiments were carried out on *jowar* in Bundi, Rajasthan, in 1960, by the ISI as well as the ICAR approaches with a view to resolving the divergence between the production estimates based on the two approaches. Even today, I do not think the main reason for the divergence was the small size and the circular shape of cuts in the ISI approach.

Second, I was invited to tabulate and analyze the time record data available from NSS Schedule 4.0 (Investigator's Time Record). Along with Shri Arun Kanti Bhattacharyya, who went over to the NSSO, I did a lot of work, in the second half of the 50's, tabulating such data for a few rounds, beginning with the 8th. However, no report was published by the NSS on this subject, for obvious reasons.

Another early recollection I have is of a visit I paid to the Department of Economics of Aligarh University. Professor Dhurjati Prasad Mukherjee was then Head of the Department, and Dr. K.A. Naqvi[1], a Reader, in-charge of a centre for socio-economic surveys to be carried out by the students, which was an activity of the NSS! Mahfooz Ahmed was one of those students. I went to advise them on tabulation and analysis of some industry data collected by this centre.

### Activities During the 60's and 70's

I joined the Planning Division of ISI, Calcutta, in 1960 and began to analyze NSS data, mainly consumer expenditure data, for various purposes. As the following account will show, I have also been interested in the methodology of level-of-living surveys. In 1961, N.S. Iyengar and I wrote a mimeographed working paper on the inward shifts in the Lorenz curve of PCE (per capita household expenditure) in rural and urban India during the 50's. This paper was widely quoted. In fact, it influenced the draft report of the Mahalanobis Committee on Distribution of Income and Wealth set up by the Government of India in 1960. Professor Mahalanobis did not approve of the draft and in 1962 began extensive studies based on NSS consumer expenditure data collected between 1951 and 1961. For several years, I assisted him, along with Moni Mukherjee, J. Roy, N.S. Iyengar and others. I learned a great deal from the Professor. He proved that our finding that the

Lorenz curve of PCE had shifted inward was very likely an illusion, caused by differential movements in consumer prices over time, adversely affecting the poor. Part II of the Mahalanobis Committee Report on Levels of Living, published in 1969, is methodologically highly interesting.

During the 60's, I did some work on estimation of engel elasticities from NSS budget data for any particular round. In one study, however, I made joint use of such data from several rounds to fit the linear expenditure system with a view to estimating engel elasticities as well as price elasticities. Later researches in ISI and elsewhere have gone ahead and made such early studies look naive and inadequate[2].

Around 1970, G.S. Chatterjee and I did some studies based on the NSS 18th round (1963-64) data to examine the variation in level of living measured by average PCE (i) between rural India and urban India, and (ii) across rural areas of different States. For this purpose, we had to estimate indices of consumer price diffrentials (i) between rural and urban areas of the country as a whole and (ii) across different States in the rural sector. Such indices were estimated from NSS budget data. These were essential for carrying out the comparisons of average PCE across regions at constant prices. Such work has also been done by others and by ourselves using NSS data of other rounds, like the 28th (1973-74).

A novel kind of sample survey was carried out in 1963-64 by Tares Maitra, Banamali De and myself in West Bengal, at the suggestion of Professor Moni Mukherjee. We wanted to study the distribution of benefits of public education and health services over the ten decile groups of the population based on a ranking of households and persons in ascending order of PCE. We were very happy later on when similar enquiries on social consumption were started on an all-India scale by the NSSO in the 80's. Data collected through such enquiries are essential for assessing how far the distribution of benefits from public services is egalitarian.

### Activities in the 80's and 90's

I became a member of the Governing Council of the NSSO around the year 1980 and served on many Working Groups formed by the Council. Contrary to criticisms often made, the NSS has not lost its dynamism. Every now and then, it has been taking up enquiries on new topics like social consumption, handicapped persons, aged persons, conditions in tribal areas - to cater to the changing needs of Indian society. The sampling design has also been continually adapted to the changing topics and environment. Computerisation has helped in expediting the publication of results and in dissemination of data. I have been deeply interested in the analysis of NSS data on level of living, poverty and inequality, in disparities in these matters across regions, socio-economic classes, etc., and in inter-temporal changes in these. I may mention three of the studies with which I have been associated.

First, four of us belonging to ISI, Calcutta, carried out studies on time trends in poverty and inequality in rural India using NSS budget data from 1952-53 to 1983 (*vide* Bhattacharya, Coondoo, Maiti and Mukherjee : *Poverty, Inequality and Prices in Rural India*. Sage Publication, 1991). This study included an econometric model which explained the observed variation in incidence of poverty over this period.

Around 1980, economic anthropologists were criticizing the finding based on NSS budget data that living standards in rural India had stagnated over time. They emphasized that they had revisited some villages after the lapse of many years and observed clear improvements in living

---

[2] I may mention that in the 90's, I did some work with Suchismita Ghose on the effect of reference period used in the consumer expenditure enquiry on engel elasticities estimated from budget data. The elasticities depend on whether data relate to the last month or to the last year.

conditions of the people. Provoked by this, Manabendu Chattopadhyay and I wrote a paper showing that judging by NSS data, there had been little improvement in household consumption up to about 1980, but there had been distinct improvements in housing conditions, mortality rates, infra-structural facilities in the village, etc. There was, in fact, broad agreement between the results of anthropological studies and the statistical surveys [see P.K. Bardhan (ed): *Conversations between Economists and Anthropologists*. Oxford University Press, 1989]

Stimulated by this dialogue with anthropologists, Ashok Rudra, Manabendu Chattopadhyay and I carried out a sample survey in 1985 in three districts of West Bengal, where we revisited villages and households covered in the NSS 27th and 28th rounds for enquiries on consumer expenditure and housing conditions after an interval of 12 or 13 years. Many aspects of level of living were examined in this re-survey. Here also it was found that standards of household consumption had not changed much over time, but there had been clear improvements in housing conditions, in household stocks of durables and semi-durables, and in infra-structural facilities in the villages. In my opinion, this re-survey of NSS villages and households pointed to directions in which the NSS methodology might be re-oriented. Throughout my career, I have made use of NSS samples of villages, households, etc., in conducting fresh enquiries, to save on time for sample selection and listing. I am happy to record that in this re-survey of NSS villages and households in West Bengal, the NSSO permitted us to use the household lists of 27th and 28th rounds and more, to use NSS 27th and 28th rounds' household data for our study on changes over time.

I was associated with the designing of some methodological studies carried out by the NSSO. One, relating to the choice of reference period in the consumer expenditure enquiry, did not yield any results; but another, on how to collect income data through household surveys, did. The Governing Council of the NSSO kindly agreed to hand over the data of this second study to us in ISI, and A.K. Adhikari and myself from ISI along with P. Chowdhury of the NSSO, produced the report in two volumes on this pilot study on Income, Consumption and Savings. It seemed to us that another pilot study on the same topic was needed for evolving the final methodology of household income surveys in India.

**Concluding Observations**

I wish to make several observations on the level of living surveys in the NSS.

First, the NSS should try to make fuller use of State Sample data so that estimates can be thrown up for regions smaller than the States.

Second, the NSS should try to make use of panel data. Changes over time are better measured and better understood if panel data are available.

Third, and most important, NSS should try to use *shorter questionnaires* so that no household interview takes more than 60 to 90 minutes. This is because respondents have become too busy and cynical about sample surveys carried out by Government or other bodies. I strongly recommend the idea which has already been mooted that for the consumer expenditure enquiry, we canvass two different questionnaires in two half-samples of households : (i) a detailed Food Schedule followed by an abridged Non-food Schedule and (ii) an abridged Food Schedule followed by a detailed Non-food Schedule. Innovative estimation procedures will have to be evolved to ensure comparability of estimates over time. Of course, any change in the methodology of data collection entails some risks; but one cannot continue with existing methodology any longer - that would be simply unethical!

# STATISTICS IN INDIA: PERSPECTIVES FOR CHANGE WITH TIME

## Ranajit Chakraborty[*]

At the time of the Golden Jubilee of the nation's independence, the Department of Statistics of the Government of India can rightfully claim to have inherited a rich tradition. Although centuries back the Mughals practised collection of statistical data for the purpose of governance, it was not until the country's independence that the government's data collection activities were closely coordinated with innovative research in the subject as well. It was possible because of the farsightedness of Professor P.C. Mahalanobis, who saw statistics not as a mere set of numbers, but as the "key technology". His vision, as pointed out by another statistician, Sir Ronald Fisher, was not simply to collect statistical data for rectifying ignorance regarding the people that a government serves; his vision was to use this "new technology for increasing the efficiency of human efforts in the widest sense" (Rao *et al.*, 1963).

At the invitation of the first Prime Minister of the country, Professor Mahalanobis led and directed the statistical wing of the Government. Earlier he had founded an institute where he initiated research and training in statistics. Thus, the early history of statistical development in India is almost equivalent to the achievements of Professor Mahalanobis and the history of his institute, the Indian Statistical Institute. His influence on collection of governmental statistics was also dictated by the fact that the two- and four-sector econometric models developed by him (Mahalanobis, 1961) were the premises under which the country's initial five year economic improvement plans were formulated. A full narration of this early history is beyond the scope of this commentary. (An excellent account of this may be found in the essay that the editors wrote on the subject of scientific contributions of Professor Mahalanobis; Rao *et al.*, 1963.)

In the paragraphs below I present, instead, some ideas on the directions in which future statistical developments can take place in India. I admit that not all of these ideas are new and some of them may not be universally accepted. However, with my training experience from India, and subsequently, as a practising statistician in which I often use governmental statistical data to formulate research problems of societal relevance, I feel that more detailed statistical data are not only desired, but also essential. Perhaps, with changing times even a shift of emphasis on new developments may be required, and I am confident that the Department of Statistics is capable of making these changes.

## 1. Missions of the Department of Statistics

At present, the official statistical system of India mainly relies on the two wings of the Department of Statistics, the National Sample Survey Organization (NSSO) and the Central Statistical Organization (CSO). The Department is also assisted by its own Computer Centre (CC) and its cadre of officers (the Indian Statistical Service) handling and analyzing data in a multitude of disciplines for the purposes of various ministries of the government.

The NSSO was originally established in 1950 as the Directorate of National Sample Survey (NSS) as a permanent sample survey organization with the mission that it will collect data to assist

---

[*] Allen King Professor, Human Genetics Center, University of Texas, School of Public Health, Houston, Texas, USA

in socio-economic planning and policy-making. The first round of work of the NSS was conducted during 1950-1951 covering rural India, under the direction of Professor Mahalanobis, in which collection of data on household income and consumption received the main emphasis. Since then, the mission of the NSSO has evolved into gathering relevant data on employment, material production and trade, crop area and yield, human resources, housing conditions as well as debt and investment and prices of commodities. Thus, the NSSO has become an institution by itself, supplying the Government the mechanism for collecting and analyzing data in every sphere of governmental activities.

The sister organization, called the Central Statistical Organization (CSO), was created in 1951 with its mission to coordinate the statistical activities of the country. The CSO's reports present some of the important official statistics of the country, with the ultimate goal of dissemination of statistics on all aspects of national life.

The ISS, instituted as the national statistical expertise system, came into being in 1961, whereby designated trained experts are to carry out the missions of the CSO and NSSO by providing relevant statistical data to the administrators. Monitoring the economic and material growth of the country, through analysis of a wealth of statistical information is also a responsibility of the ISS officers.

The Computer Centre (CC) of the Department of Statistics, the youngest to appear on the scene (in 1967), plays a pivotal rôle in conducting analysis of data gathered by the NSSO and CSO. It also maintains their databases and imparts intensive training for data processing and systems analysis.

## 2. The Future of the Department of Statistics

In highlighting the 1997 annual national budget, on its webpage (www.indiagov.org/economy.menu.html), the Government of India proclaimed that changes were taking place on the economic and political fronts within the country. These changes, prompted by a new spirit of economic freedom, are geared to erase the common notion that this seventh largest (by geographical area) and second most populous country of the world is also the one of unrealized potential. The fact that India's time-tested institutions offer an environment where long-term investments can be trusted with security, where the presence of a vibrant free press ensures public accountability, where the judiciary can and does overrule governmental follies and wrong-doings, where the highly skilled managerial, technical manpower and intellectual infrastructure can offer the backbone of competitive endeavours, etc., are among the rationales that are cited as roots for launching the free economy movement. The statistics compiled by the Department of Statistics underlie the basis of these claims. Further, such statistical data, collected during the last fifty years, document that India has no fundamental conflict between its political and economic systems. This is also aided by the size of the middle class population, that exceeds the total population of the USA or that of the entire European Union, providing the country with a definite cutting edge in global competition.

These changes are certainly welcome and the initial indication of their success has also been evident since 1991. However, with these changes the responsibility of the Department has also to change, because without the government-sponsored data collection on projects of international collaboration, the intellectual property and commercialization rights of joint ventures may not be ultimately beneficial for the country.

An example of this may be given from the biotechnology area where India has made considerable progress and there appears to be a substantial international interest. Propagation of plants by tissue-culture, a time-tested, well-established and widely used technology in India, is now almost entirely biotechnology-based in which foreign participants have shown interest for its application in floriculture industry. It is estimated that with the variety, uniqueness and abundance of the country's flora, India can earn 25 to 30 times as much of foreign exchange from exports in this industry as from exporting other agricultural products. However, if this industry is developed, it is still unclear how the intellectual rights of the Indian floral heritage will be preserved. While the international patent laws are complex and their jurisdiction is outside the normal domain of the Department of Statistics, from the recent episode of the *Basmati* rice patent dispute in the USA, it can be learned that the documented statistics of uniqueness of Indian resources can serve the purpose easier for preserving the Indian interest on such rights.

In the area of human genome research as well, similar considerations may apply. India, with its unique population diversity, has attracted anthropologists and geneticists for over a century. Because of technical limitations, many of the current human geneticists in India have been collaborating with foreign experts. Often such joint research ventures lead to transporting human DNA samples out of the country whose documentation after the conclusion of the initial study is not known. Since such samples can be preserved over a long period of time with the current technology, the question remains as to who has the right to any future intellectual information on these samples. The Department of Statistics with the help of the Department of Science and Technology (DST) should formulate mandatory means of statistical inventory of such research.

While the above examples perhaps constitute areas of minimal national priority, on the front of national health as well, statistical data collection methods may need a substantial shift of emphasis. By and large, ill-health in India is mostly due to malnutrition and early growth and development problems. With average life-expectancy at birth of a little over 60 years for males and 61 years for females (1992-93 data), diseases of the aged are still not among the most urgent areas of interest. Nevertheless, national as well as State-wise data on diseases and disabilities should be collected and maintained by ICDA codes, and both governmental as well as non-governmental agencies should report such data for inclusion in the Department of Statistics databases. Such data, collected over a period of time, can offer guidance with respect to the changing disease profile of the country. Further, for formulating any future health care system at the national level, such data can make a valuable contribution.

A matter of somewhat urgent need in this area is a more detailed documentation of the AIDS epidemic in India. It is often said that with the current increasing rate of AIDS infection, India will soon have the largest number of AIDS patients in the world. Yet, even the experts in the field agree that HIV statistics in India are incomplete and can be misleading as well (Bharat, 1996). While some attempts are being made to rectify this drawback, the Department of Statistics, with its statistical expertise, can ensure the legitimacy of the estimate of rate of infection and its breakdown for devising an efficient modality of preventing any further pandemic increase of this deadly disease in India.

Finally, unfortunately for India, administrative corruption is a hindering aspect of the country's national growth. While the free press and the independent judicial system of India often can and do bring such corruption to justice, this causes a great loss of national efficiency and expenditure of national resources. Perhaps, the official statistical system may be able to build means of accountability of public officials through which evidence of corruption at its early stage may be detected and brought to attention. This may require an innovative strategy of monitoring and

analysis of expenditure on governmental projects where the individual privacy rights of a citizen should be acknowledged and honoured, yet misuse and misappropriation of authority can be rectified and controlled at their earliest inception.

## 3. Concluding Remarks

With these comments notwithstanding, I would reiterate that the Department of Statistics, in its first fifty years of service, has not only served the country well, it also has trained a number of eminent statisticians many of whom are now helping other nations as well for the betterment of their economic prosperity. All this has been due to the vision of the founding father of this Department, and it has been through the diligent work of this Department, that India has now a pivotal place in the world map of statistics. Thus I am confident that should the above comments have any significance in the context of the future of the country, this Department is certainly capable of addressing the relevant issues.

---

## References

Bharat, S. (1996). *Facing the Challenge: Household and Community Response to HIV/AIDS* in *Mumbai, India*. Global Programme on AIDS. World Health Organization, Geneva, Switzerland.

Mahalanobis, P.C. (1961). *Talks on Planning* (Collected addresses and broadcasts on various aspects of planning in India). Asia Publishing House and Statistical Publishing Society, Calcutta.

Rao, C.R., Lahiri, D.B., Pant, P., Nair, K.R. and Shrikhande, S.S. (1963). *Contributions to Statistics*. Pergamon Press, Oxford and Statistical Publishing Society, Calcutta.

# STATISTICS, DEMOCRACY AND DEVELOPMENT

## Carlos M. Jarque[*]

## 1. Introduction[1]

Knowledge knows no distances : it travels unbounded, impelled by its own vitality. When I was a graduate student in England, and later in Australia, ideas and papers on statistics by P.C. Mahalanobis greatly helped me in the study of various matters of cluster analysis that concerned me at that time.

First, I would like to express my respectful homage to the memory of Professor Mahalanobis, founder of the Indian Statistical Institute and Honorary President of the ISI (International Statistical Institute) since 1957, for his contributions to statistical science in general, to the methodology of planning, and to the direct and indirect training of many colleagues (see Rao (1978) for a detailed biographical reference).

In this article I shall concentrate, primarily, on the analysis of several aspects relating to Official Statistics and to its contributions to democracy and socio-economic development. From the start, it is worth pointing out that it is not the intention to present original methodological developments, nor to carry out modelling exercises to establish causality between statistical development, democracy and welfare. In fact, the objective is to make some reflections, that perhaps should be made more often, on the purpose of Official Statistics, the resources allocated to it and the benefits derived from this investment.

As it is known, the consistent reduction of financial resources in some regions, both in Official Statistics and in areas of statistical research and teaching, has become a point of increasing concern to the statistical profession. See Bonnen (1983) and references cited therein and also ISI (1987) and ISI (1989). We intend here to emphasize the *rôle* of statisticians, by highlighting its useful work and positive contributions in modern society. Some of the arguments and experiences presented refer principally to Official Statistics and to the developing world. Yet, we feel that they are not exclusive of these.

In the first part, and to give a brief historical perspective, reference will be made to the origins of Official Statistics in Hispanic America. Then reflections will be carried out on the current world situation and its influence on the *rôle* of Statisticians. To follow, for the case of Latin American country, Mexico, the impact that this context has had on the structure of the statistical system and the resources allocated to this will be mentioned. Here a series of recent actions which have allowed the improvement of statistical information in the short term will be illustrated. Finally, considerations will be presented on the contributions of statistics to democracy and development.

## 2. Official Statistics in Latin America

Systematized information on the physical environment and its inhabitants helps to identify the needs of people and to measure the potential of any society. Information is at the very heart of our ability to act effectively and successfully.

It is no wonder that nations give a high priority to knowledge both of themselves and of their surroundings. Thanks to sufficient, timely and high-quality information, they can reduce the disadvantages of uncertainty and of a life dominated by chance. As we know, the statistical count, method and technique yield precious information with these valuable features. So, statistical activity plays a strategic *rôle* in the development of peoples and nations.

As in other parts of the world, in Latin America this has long been recognized since remote times. There are many proofs from the region's ancient cultures which show their interest in statistical and geographical information as a matter of state. We may recall, for instance, that the great civilizations which evolved on the American continent a millennium ago, such as the Aztecs, the Maya, Olmecs and Incas, were already carrying out systematic statistical studies. This is evidenced in Codices and other historical documents which are now part of humanity's cultural heritage: moreover, near Mexico City there is a place with stone hills known as "Nepohualco" in Nahuatl or "Contadero" in Spanish meaning "Counting Place" (see Flores (1958)). Existing hieroglyphs admit that the hills were built by the Chichimecas around the year 1000 A.C., with the stones deposited by each inhabitant of the region with the objective of doing a statistical count. Indeed, some consider this place to be a census data bank of the eleventh century.

These ancestral cultures were also interested in cartography. Thus, they gathered information about the territory they inhabited, drawing up numerous geographical charts and maps. Their interest in relating socio-economic information to small geographical areas, is registered in many paintings and codices that constitute valuable testimony of the importance given to these activities.

With the arrival in America of Christopher Columbus in the year 1492, these American civilizations came into contact with the European cultures of the time. This event, called the encounter of two worlds, led to an intense mutual influence on ways of being and of understanding the universe. In fact, knowledge, culture, the arts and the sciences were all enriched. As part of the five-hundredth anniversary of this encounter, many analyses were done of the complex and brilliant synthesis which emerged, in many fields, from that coming together of civilizations both different and simultaneously valuable.

The statistical methods used by American Indians were also enriched in this flourishing. Indeed, during the post-Columbian period, their procedures were complemented by those of the European authorities carrying out work in the compilation of statistics. Among others, we could cite the influence of the Spanish, English, French and Portuguese schools, which met with American colleagues already in the sixteenth century. Thus were born new statistical methods and guidelines, with characteristics and procedures specific to the region.

Later, as the countries began evolving from mainly agricultural societies towards increasingly industrial ones, at the end of the nineteenth century, governments throughout the newly-independent American nations became involved in setting up central offices for official statistics. Their *rôle* would be, and was, to compile statistical information systematically as a basis for decision-making. It is thus no coincidence that many of the central statistical offices in the hemisphere were founded towards the end of the nineteenth century or the beginning of the twentieth.

A similar development may be observed in many other countries of different latitudes, in which the central statistical offices emerged around this period.

Now, a thousand years after the first systematic efforts to generate statistical information in Latin America, statistical information has become a crucial factor in the political, economic and social development of nations.

## 3. The World Transformation

In recent months we have seen immense changes in the architecture of the world. Countries of all latitudes have entered into processes of transformation in order to fulfil the needs of present-day life.

Millenarian societies, centuries-old nations, recent independent states, union groups and political parties of all kinds are in a process of modernization. Rhythms, strategies and forms may vary: in some cases, interest is centered on the extension of democratic processes; in others, on the reform of trade or on the creation of new economic markets; in still others, mechanisms are painstakingly refined so as to allow for better social developments.

This modernization phenomenon is generating a more interdependent and competitive world, before which no one can, or should, remain indifferent. It is a context in which statisticians have an active *rôle* to play, as they had in other periods of marked historical importance.

In this world situation, no nation has its future guaranteed, and to fall behind is to opt for marginalization. That is why each country must channel efforts and define strategies, according to its own circumstances, so as to take full advantage of its potential.

As concerns Mexico, the modernization strategy has been inscribed in the National Development Plans, which are drawn through extensive popular consultations (see SPP (1989)).

The main premises of these are political, economic and social reforms aimed at establishing the necessary and sufficient actions so as to eliminate the barriers that stand between modernization and present conditions that are obsolete due to internal demands and to the new world context. To support this process, an analysis was made of the changes required in the country's institutions, in particular in public agencies and, specially, in the statistical offices whose efficient operation was identified as an indispensable factor to pretend and to achieve modernity.

## 4. Statistical Activity in Mexico

In Mexico, Official Statistics are coordinated by the recently created National Institute of Statistics, Geography and Informatics (INEGI). However, its constituent parts are not all this new : the Central Bureau of Statistics, for instance, was founded in 1882, more than one hundred years ago.

The Institute currently includes a Presidency, six General Bureaux in the Central Offices and ten Regional Divisions. This allows it to service a population exceeding 96 million inhabitants, over a territory of two million square kilometers.

INEGI's Central Bureau of Statistics is Mexico's General Statistical Office. Its activities, among others, include Population and Housing Census; Agricultural, Livestock and Forestry Census, and numerous monthly and quarterly surveys, in households, businesses and firms. This Bureau also uses administrative records to generate, among many others, the vital statistics, balance of trade figures and indicators of cultural activity. It is also in charge of the Economic Censuses, the National Accounting System, the quarterly and regional Gross Domestic Product figures, and other macroeconomic indicators.

A second component of the INEGI is the Central Bureau of Geography. It uses aerial photography, teledetection, remote sensing and research, which together with ample field activities allow the carrying out of geographical studies of Mexico. For example, this Bureau quantifies areas of idle land with agricultural potential, forest areas lost to fire and phenomena affecting our

ecosystems. It also provides the official geodesic positioning service and carries out a large activity to elaborate the country's basic and thematic cartography.

To illustrate the amount of information this signifies, it may be mentioned that the basic or topographical map of Mexico, scale 1 to 50,000, consists of 2,440 sheets, which form a rectangle of 60 meters from east to west, and 45 meters from north to south. This presents all the orographical characteristics, roads, highways, paths, human settlements and other topographical aspects. It is very useful material, which requires constant updating to preserve its value. The INEGI also produces thematic maps covering hydrological, geological and edaphological aspects, as well as soil use and the potential use of the territory, among other things. In 1992, new cartographic activities were given to the INEGI, in a cadastre programme for land titling of 102 million hectares (half of the country).

It is worth noting that the incorporation of both Mexico's Central Statistical Office and the Bureau of Geographical Studies in one institution follows a millenarian tradition, offering, among other things, great advantages in terms of ease and precision in referencing and accessing socio-economic information in its spatial dimension.

The INEGI also includes the Government's Central Bureau of Informatics, which, aside from processing information, establishes the Federal Government's computer services development policy, promotes the use of computers in State and municipal governments, and sets guidelines for the public sector's purchases of computer equipment and services.

These are three of the six General Bureaux of the INEGI.

A fourth Bureau is that of diffusion. It includes the printing facilities which produce the 900 regular titles published by the INEGI, administers the network of 600 service and reference centers, heads the public relations campaign and attends to journalists. A fifth General Bureau is in charge of legal, auditing and training related matters, and a sixth General Bureau is in charge of administration and international relations.

This brief description of the Institute's central structure, which is complemented by its ten regional offices, in charge of execution and field operation of projects, gives a general view of the Official Statistical work in Mexico. But of more significance is the reflection upon the challenges that this, as other similar institutions, face in the early nineties within a globalized and interdependent world, and within a nation fully committed to modernization.

## 5. Institutional Modernization

The INEGI has established a strategy to fulfil the demand for information which would grow, very quickly, as Mexico implements a wide ranging modernization reform.

Before defining the programme's technical characteristics, the strategy sought to obtain, for statistical information, a high political priority, which would have to translate into a more solid budgetary support.

Despite the existence of a strict economic stabilization programme which led to budget cuts in many government agencies and the restructuring of the public sector (going from more than 600 to less than 300 public companies in only two years), the INEGI's budget grew significantly in real terms. Actually, we might say that, as part of the strategy for the Reform of the State, public resources were re-allocated from non-priority areas (such as certain firms which were owned by the government) to consolidating strategic ones, like the production and diffusion of statistical information.

In 1983, the public sector in Mexico generated 25 percent of GDP; by 1991, after the reduction of administrative areas and the privatization of the banks, the telephone company, airlines and other assets, this percentage was redimensioned to 16 percent. The areas of activity of the State were thus re-defined and, without neglecting other strategic functions, it designated statistical information as a high-priority area for public action. With the support of the country's highest authorities, the INEGI was able to consolidate current activities and to undertake new ones, all aimed at promoting institutional efficiency.

Thus, the INEGI defined and is implementing its own modernization programme, so as to attain its fullest potential to support a political, social and economic reform in Mexico. The INEGI's programme, granted this political priority and support, includes various aspects which I would like to discuss briefly in order to illustrate some of the many options that may effectively increase the *rôle* of statisticians.

**Decentralization** A National Statistical System must produce information with a disaggregate geographical basis. Given the size of Mexico's territory, special efforts were required to move from a geographically centralized system to an alternative allowing more contact with information sources and users at the regional and local levels. The modernization programme thus included decentralization through the consolidation of the institute's ten regional divisions. These reflect the INEGI's central organization, with areas for statistics, geography, and computer services, under the supervision of a Regional Director who reports to the President of the Institute, and who represents the INEGI before State governments. The work done by these regional offices is so important that they have become a decisive factor in the planning carried out by the local governments. In the last two years, this regional structure has been consolidated and enriched by the creation, and now full operation, of 32 State offices and 1,700 municipal offices.

The institute's decentralization also included the transfer of its central offices from Mexico City to Aguascalientes, 500 km. to the north, in the country's geographical centre. This process began after the earthquake of 1985, which caused significant losses to our office space in Mexico City. The transfer of the INEGI's central offices, as part of a general relocation, away from the capital, of government agencies, implied building offices, workshops and laboratories. Similarly, almost 3,500 housing units were built for the relocated personnel and their families, and urban services had to be extended to supply the needs of this new population in Aguascalientes.

**Infrastructure** A second aspect of modernization includes the equipment to carry out the work. As is well known, the scientific and technological revolution produces ever more powerful equipment at incredible speed, to set up computer networks, teleprocessing systems and instruments to interpret satellite images. This led to the undertaking of a diagnosis of our current equipment, in order to progress towards more advanced and proven technologies.

We thus identified our need for more equipment: As a result, 6,000 new personal computers were put into operation with specialized network; the ten regional computer centres were linked via satellite; equipment was acquired to digitize cartography and to create geostatistical databases, and new photographic equipment was installed in the INEGI's air fleet consisting of Lear jets, Cessnas and helicopters.

As is well known, there is a period of learning and adjustment to new technologies which in our case lasted about two years for the computer networks. However, these changes have already raised the INEGI's productivity.

**Methodologies** A third aspect of the programme was a review of methodology, carried out by each area through discussion and analysis with the main groups of users. This allowed the thorough examination of the methods used and the further identifying of areas not yet covered by Official Statistics. This led to new projects to support and enrich the national system. Among these

numerous projects, we could cite a large study of the informal economy, the development of ecological accounts, the estimation of delinquency not reported to the judicial system, and an extensive study together with the Ministry for Public Education to evaluate the terminal efficiency and quality of the country's primary and secondary education.

It is worth noting that, in the task of methodological revision, as regards both traditional and new projects, the research done in academic institutions has played an important *rôle*. That is why very close contacts have been cultivated with the scientific and academic sector.

The actions undertaken by the INEGI in this respect include conducting many workshops and seminars on the projects to obtain feedback from specialized users of information. Also, meetings are organized with other producers in order to share experiences. Of special relevance is the INEGI's support in re-establishing the Mexican Statistical Society in 1989 - an event of great operational and historical significance. It now includes over 300 professionals in the field and represents an important advance in the development of statistics in our country and, we believe, also in Latin America.

**Training**  The training of its human resources is another component in the INEGI's modernization process. Because they are the principal asset, they must be trained in an on-going fashion, not only to follow advances in their work infrastructure and new methodologies, but also to stimulate their own full development.

With this goal in mind, a Training Programme was established which includes, among other things, four diplomas of a 14 month duration. These began in 1990 and are now recognized by the Ministry of Public Education as short professional careers. To implement this programme, there are classrooms, a library, a computer centre, as well as INEGI-owned apartments to house 150 scholars. These facilities are being further enriched by a training centre, built specifically for that purpose. The training programme also has an international component.

**Coordination**  The INEGI's current work goes beyond mere data production; it is also responsible for the coordination for establishing a National Information System.

To this end, it has been important to carry out joint projects with all other sectors and with the States of the Federation, so that each one can generate information according to its own needs, but within a uniform conceptual framework.

**New Products**  The ample social changes and technological advances are generating a broader, more sophisticated range of information users, and it is important to identify their needs and to design products to satisfy them adequately.

With this objective, and thanks to budgetary savings due to increased productivity, the INEGI has added to its structure an interdisciplinary area responsible, among other things, for market research, technological monitoring and the development of new and better products.

In developing this wide range of new products, marketed as diskettes, compact discs and databases, the need has been recognized for supplying consumers with user-friendly analytic software. Part of this has been done at the INEGI and delivered to users free of charge, in order to promote the computerized use of information. However, with some other products, systems developed by private firms have been included allowing for more specialized use.

**Broader Statistical Culture**  All the aspects of the modernization programme that have been mentioned so far, such as decentralization, work infrastructure, methodological revision, training, coordination and new products, are important in improving the information services. However, these

efforts would have limited impact and would be insufficient as long as segments of the population do not know about the information generated or are unable to use it in their own decision-making.

Extensive campaigns were conducted to promote participation by citizens in various statistical projects. As an example of its breadth, it may be noted that the phase of the campaign for the 1990 Population and Housing Census alone used half of the official mass media radio and broadcast time which, by law, has to be assigned for government programmes.

After this wide campaign a survey was carried out, which showed a substantial increase in knowledge about the institution. However, it was also found that, despite advances, a part of the population did not know the specific figures of the results obtained. In fact, the main conclusion of these studies was that it is of utmost importance to advance towards a greater culture in the use of statistical information.

It would seem that, despite the application of twenty-first-century technology to produce information by agencies, some segments still use it but rarely and with obsolete methods. In the face of this, a conviction has been reaffirmed that perhaps the statistician's duty should not be limited only to producing with accuracy relevant data; but he should also assume some of the responsibilities for later phases, such as ample diffusion in order to promote its intensive use, giving rise to wider statistical marketing.

It is important to note that, traditionally, almost all budget resources are used to produce information, with little left over for its dissemination. That is why the INEGI has undertaken actions not only to expand its resources, but also to change its budget structure, giving a well-deserved priority to dissemination.

## 6. Statistics, Democracy and Development

These are some of the measures recently taken to modernize the INEGI, in order to advance in one of Mexico's priority areas: an efficient Public Service for Statistical and Geographical Information. Of course, most of these measures are not new to our community, and others require on-going and permanent efforts. Yet, what is interesting is that these actions, defined as parts of an integrated programme, have been simultaneously implemented with greater political and financial support.

Certainly, as in the past, many challenges remain. At the moment, some results are clear and give us good hope for further progress. These advances may lead us to the following questions:

Why has statistics been assigned one of the activities of higher political priority ?
What is, or should be, the contribution of statistics to the noble goals of democracy and development ?

I believe these questions apply to all countries but I will intend to give answers for the particular case of Mexico.

In referring to the government programmes, we mentioned the National Development Plans. These set the guidelines for the progress of Mexico. Their strategy touches upon three basic objectives: firstly, extension of democratic participation and practices; secondly, economic recovery and growth with price stability; and thirdly, improvement in the living standards of the people.

What is the *rôle* of the statistical projects in the advances of these Agreements which, in turn, are promoting Mexico's modernization ?

**Democratic Participation and Practices**

Let us examine the first objective, aimed at extending democratic practices.

**(I) Statistics and the Right to Information**

The efficient dissemination of statistical information helps in bringing about a better-informed society, with greater awareness and, thus, greater participation. In Mexico, as in many other countries, the right to *Information* is a constitutional mandate. The fulfilment of this requires disseminating objective, timely and clearly intelligible information. In this way, each individual may form his own, reasoned opinion - which, in our view, is one of the necessary conditions for an effective democracy.

Society requires and demands today ample information including, of course, accurate statistics in order to form well-grounded judgments regarding the performance of its leaders and institutions and about the country's overall progress. Only thus can people make decisions and choose among the various options available, in the matters that concern the life of the nation.

In the pursuit of this goal, statistical offices and statisticians have an inescapable public responsibility: to carry out their work within the strictest professional ethics (see ISI (1986)), independent of any interests other than those of their competence. This is the only way by which they can achieve credibility, and society's trust, in the information generated. This responsibility must be assumed in order to fulfil the right to information of the current and of the future generations.

**(II) Population Statistics and the Congress**

In addition, statistical endeavours support democratic practices by giving updated population figures. These make possible, for instance, the demarcation of districts for the representation of the Federal and Local Congresses. Here too, there is a clear commitment to use objective and scientific procedures in obtaining the population figures, to ensure that popular representation is fairly implemented.

**(III) Statistics and the Electoral Process**

Other contributions of statistics to democracy may be noted. For example, in Mexico, one of the demands made by all political parties for the congressional elections of 1991 was to establish a new electoral roll. The request was that it be done from scratch -mainly because of the intense migration processes in the country during recent decades. The new electoral roll was deemed indispensable for the country's full democratic process by all sectors of society.

Since the INEGI was to carry out the 1990 Population and Housing Census (hiring close to 700,000 people for this task), the political parties unanimously agreed that the Institute should simultaneously generate the voters' roll.

However, as we know, it is a delicate matter to mix Official Statistical work with electoral processes. With good reason, this topic has been widely debated in various fora. Because of the clear importance of both tasks, they must be carried out with full transparency and respect for ethical principles, and must comply with the legislation governing the agencies involved. In Mexico, the law establishes a complete separation between the Official Statistical Information activities, and the electoral, fiscal and judicial systems. It also guarantees the confidentiality of all individual data provided by institutions, firms or persons.

The government was thus confronted with two goals whose simultaneous fulfilment required careful attention and negotiations: on the one hand, it was necessary that the juridical framework be respected, keeping the INEGI separate from electoral matters; on the other, that the best technical means be used to fulfil, in a very short time, society's demand for an updated electoral roll.

These negotiations led to a result where in March, 1990, the INEGI carried out the population census as stipulated by the Statistical Information Legislation, completely apart from electoral matters. Afterward, the technical group which planned and carried out the census acted as a technical unit to the electoral authorities and political parties so as to undertake an additional project, eight months later; its purpose was to develop the new voters' roll, using questionnaires different from those of the Statistical System. To this end, the INEGI provided, among other forms of support, cartographic material derived from the census, field work organization principles, personnel-training technologies and administrative procedures. Also, in order to take advantage of existing experience the electoral authorities recruited people who no longer worked for the INEGI but who had participated in the census.

Thanks to these resources and a broad citizen participation, the new electoral roll was developed between January and June of 1991. Obviously, this arrangement implied additional funding over and above that devoted to the Population Census. But the separation between the Official Statistical and Electoral systems was maintained, and an important social need was fulfilled. The new roll was used, as programmed, in the elections of August 18, 1991, within a framework of greater efficiency, and with one of the highest rates of voter participation in Mexican history. In addition, in the electoral legislation it has been stipulated that the INEGI is to permanently advise, in its areas of expertise, the work of a newly-created Federal Electoral Institute.

## (IV) Conclusion

To sum up the statistical projects in Mexico have become a necessary infrastructure to implement the full right to information, to help ensure the population's fair representation in the legislative branch, and to make more effective the mechanisms by which one can exercise the right to vote. It can thus be appreciated how statistical activity is not only a technical endeavour, but also a task of the greatest political significance. In other words, there is a positive interrelation between statistical development and the democratic life in a country.

### Economic Growth with Price Stability

Regarding the contribution of statistics to the second agreement, for economic recovery with price stability, it may be observed that the strategies allowing for greater production of goods and services, without inflation, need, as a fundamental element, accurate and timely information. The adoption of fiscal, monetary, trade, wages and financial policies, to mention only a few, require this information as a crucial input.

In Mexico, the recent strategy to cut inflation included, apart from the orthodox measures of demand restraint, a social pact to control and administer prices and wages. Therefore, the National Statistical Information System had to complement its projects in order to register possible shortages of products which might stem from such a programme. Similarly, it was necessary to supplement the measures of price indices, based upon various baskets of goods according to socio-economic levels and geographical zones. With these and other statistics, a weekly analysis was made of the evolution of various aspects of the economy to support the synchronization of the instruments of economic policy.

In addition to these requests from the public agencies, the private sector requires statistical information upon which to base its plans for investment, production and distribution. Its demand for

information has also increased in recent times in order to carry out productivity and competitiveness analyses in the face of free-trade agreements in North America. In this regard, information derived from the 1989 Economic Censuses and Establishment Surveys, for example, has played a fundamental *rôle* in studying policies impact on the sectoral and regional levels, and to measure the speed at which changes take place in an open commercial market.

Mexico's heterodox economic stabilization programme to combat inflation and recover growth has made special demands for new statistical projects in the information system. Some of these, due to their importance, have become part of the INEGI's regular activities.

## Improvement of Living Standards

The third objective concerns the improvement of living standards, with greater participation and solidarity. In this regard, social statistics covering the characteristics of jobs and unemployment, health, education and the environment, as well as public safety, poverty and housing conditions, have allowed diagnoses leading to policies to alleviate shortages, and to actions aimed at improving living conditions.

After almost a decade of economic crisis, with successive austerity programmes and low levels of public revenue, a number of social lags accumulated had to be attended to. Programmes to combat poverty were thus established, with novel systems whereby funding is provided by the Federal and State governments, and also by the beneficiaries themselves, through their work or other inputs.

These programmes now represent a significant percentage of the Federal budget. To implement them, it was necessary to have information in order to detect and identify exactly where problems existed, and with what intensity. Information provided by the 1990 and the 1995 Population and Housing Censuses proved extremely useful in this regard, since they provide data on small geographical units, covering a city block or group of blocks, thus avoiding the use of aggregate indicators which carry the risk of being unable to identify specifically the human groupings most in need of attention. In this way, the government insures that resources and subsidies reach the target population.

## 7. Results of the Strategy

The progress achieved in the above-mentioned areas, for increased democracy, economic development and welfare, gives the country more encouraging prospects than it had before.

Some results of this strategy solidly justify these expectations. The advances in Mexico are important indeed, but of course we recognize the many challenges still ahead and the efforts still needed. One may only note, for example, that if the current rate of growth of the population of 2.1 percent a year is maintained, the population will double in 33 years. In other words, in about 3 decades it would be necessary to replicate the infrastructure that the country has built throughout its long history and, simultaneously, to provide a solution to demands already existing and unfulfilled.

## 8. Conclusion

Throughout history, statistical information has played an important *rôle* in the transformations undertaken by societies. Thanks to statistics, we can apprehend reality, make a better use of our resources and contribute to a better-informed society. Therefore, Statistics are crucial in consolidating democracy and in promoting development. At present, the statistical systems encompass advances which would have been unthinkable only a few years ago. Yet, given the new scenario, in which a rapidly changing world is emerging, it is necessary to redouble efforts and to give a new impulse to our activity.

Surely the work of each Statistical Office, department, or professional group, will be more effective if it is enriched by the sum of the regional and the international knowledge and talent. There is no reason to confront our challenges alone, least those that are common to us all.

Certainly, statistical activity, our profession, is methodologically interesting, technically challenging and scientifically fascinating. But I believe its truest beauty and nobility is to be found in its contribution to helping nations to shape their own destiny and to be better equipped to provide their inhabitants with justice and well-being. In my opinion, never before had statisticians a better opportunity and greater technical resources to contribute to the attainment of these elevated goals of humanity.

## References

Bonnen J.T. (1983). "Official Statistics in Troubled Times: The Changing Environment of Producers and Users"; *ISI Bulletin*, Madrid, Vol. L: 2, 836-853.

Flores R. (1958). "Historia de la Estadistica Nacional"; Sociedad Mexicana de Geografia y stadistica; V. LXXXVI, 3.

I S I (1986), "Declaration of Professional Ethics"; *International Statistical Review*, 54, 2, 227-242.

I S I (1987), Tokyo, Minutes of the General Assembly, *Bulletin of the International Statistical Institute*.

I S I (1989), Paris, Bulletin of the International Statistical Institute, Vol. LIII: 4; P. 328. Presidencia de la Republica (1995), Mexico, "Plan Nacional de Desarrollo 1995-2000"

Rao C. R. (1978). "Prasanta Mahalanobis"; *International Encyclopaedia of Statistics*; Sree Press, 571-576.

S P P (1989), "Plan Nacional de Desarrollo 1989-1994"; Mexico, Secretaria de Programacion y Presupuesto.

# SOME THOUGHTS ON INDIAN OFFICIAL STATISTICS

## S. Ray[*]

### 1. Introduction

Official statistics were being generated as a by-product of administration even during the British period. These secondary statistics were augmented with primary statistics collected through censuses and sample surveys. During the nineteen fifties, official statistical system took a concrete shape with the compilation of National Accounts, Annual Survey of Industries, agricultural statistics, prices and price indices and socio-economic statistics relating to consumption habits, employment pattern, educational status, etc. These statistics are being compiled for national planning in various fields of national development over the years. Along the pace of development, there has been a demand for a greater volume of statistics and for a quick dissemination of them. Various measures have been taken at various stages in order to bring out useful statistics in the shortest possible time to get the maximum benefit out of these statistics. These measures are continuously reviewed in order to bring about more and more improvement in them. An attempt has been made in this paper to evaluate the present status of official statistics and to suggest new areas that may come up in the near future so that effective steps can be taken to steer properly the statistical system in the country.

### 2. The Present Status

Some of the major statistics that are compiled by various departments of the Government of India are described below :

**Industrial Statistics** Industrial Statistics are collected through the Annual Survey of Industries (ASI) by the Department of Statistics by covering the census-sector as well as the sample-sector factories that are included in sections 2m(i) and 2m(ii) of Factories Act, 1948. In addition, it also covers *bidi* and cigar manufacturing establishments registered under the *Bidi* and Cigar Workers (Condition of Employment) Act, 1966, all the electricity undertakings engaged in the generation, transmission and distribution of electricity registered with the Central Electricity Authority and the activities of cold storage, water supply and some repair services. The census-sector factories are covered by census and the sample-sector ones are covered through a sample survey. The size of this sample used for survey in the sample-sector is quite large. The results are available about a year from the completion of the survey. However, attempts are being made to bring the results within 6 months from the date of completion of the survey.

**Socio-Economic Statistics** The National Sample Survey Organisation (NSSO) generally follows a 10-year programme of socio-economic surveys in which three surveys are quinquennial in nature and one is decennial in nature. The rest of the period in the 10-year programme are covered by *ad hoc* surveys of current interest. The key results are brought out within nine months after completion of the survey. Measures are being taken to further reduce the time of processing.

---

[*] Chief Executive Officer, NSSO. The views expressed in the paper are of the author and not of the organisation to which he belongs.

**Price Statistics**   Prices and price indices are presently being compiled by various Ministries. Ministry of Industry compiles and releases Wholesale Price Index (WPI), Ministry of Labour brings out monthly Consumer Price Index for Industrial Workers [CPI(IW)] and Consumer Price Index for Agricultural Labour [CPI(AL)]. Department of Statistics, Ministry of Planning & Programme Implementation, brings out monthly Consumer Price Index for urban non-manual employees [CPI(UNME)]. The time taken for release of these indices is not the same for all the series. An attempt is being made to evolve a general-purpose CPI and release the index within four weeks from the end of the month for which the index is compiled.

**Informal Sector Statistics**   The informal sector statistics are generated through sample surveys. The frame for such surveys is prepared once in 10 years through the conduct of Economic Census of non-agricultural activities spread all over the country. The estimates of different sub-sectors like manufacturing, trading, services, etc., of the informal sector are obtained by conducting follow-up surveys based on the frames prepared from the results of the economic census. The economic census will now be conducted quinquennially in order to update the frame with greater frequency.

**Agricultural Statistics**   Agricultural Statistics are compiled by the States and then co-ordinated as well as consolidated by the Ministry of Agriculture, Government of India. The Ministry of Agriculture generates advance estimates of area and yield through Timely Reporting Scheme (TRS) and General Crop Estimation Surveys (GCES), but final results are available only one year after the close of the Agriculture year. Moreover, these estimates do not include many items like horticulture, floriculture, etc., which do not fall under the category of principal crops. This gap has to be filled.

**National Accounts Statistics**   Presently, National Accounts Statistics are compiled once in a year and there is a feeling that Gross Domestic Product (GDP) is under-estimated. The reasons of under-estimation are the following:

(a) Rates and ratios for selected economic activities that are used to update GDP are outdated.

(b) There is a reasonable chance of under-reporting by industrial undertakings.

(c) Lack of identification of contributions from various segments of the informal sector which includes in itself Small Scale Industries (SSI), some registered units of ASI, Own Account Enterprises and Directory/Non-Directory Establishments.

(d) Non-reporting in respect of certain crops like horticulture, tissue culture, etc., and under-reporting in the case of meat production and other livestock products.

Necessary measures are being taken to fill these gaps. In addition, in order to meet the international requirement, GDP will also be compiled quarterly, in future.

**Population Statistics**   The Indian census is conducted once in ten years and the last one was conducted in 1991. This census is conducted by the Office of the Registrar General of India, Ministry of Home Affairs. But except for a few indicators, their results take lot of time to be released .

## 3. Proposal for the future

In order to compile the official statistics in time and to ensure quality, the following suggestions are made:

**Statistical Authority**  Department of Statistics (DoS) requires to be amply strengthened with appropriate authority so that it can do proper co-ordination and control statistical activities in order to ensure timely generation of quality statistics, quick dissemination of them, optimum utilisation of resources and avoidance of duplication. The office of the RGI may be brought under the DoS so that it can properly synchronise the population census and the economic census. This synchronisation will throw up a properly identifiable frame that can be used for follow up surveys of the informal sector. The villages/blocks may be given permanent numbers which will help computerisation. Besides, village/ block maps will be easily available for survey purposes. This authority may also take up some collaborative studies with other Departments to generate some useful indicators for social benefit. For example, drinking water is one of the necessities of human life, but its quality has not been studied throughout the country. With the Central Pollution Control Board, a collaborative study can be taken up to study the quality of water at different locations every month and with these data a water use index can be worked out.

**Proper Identification of Informal Sector**  A high percentage of GDP comes from the informal sector, which include of Small Scale Industries (SSIs), Own Account Enterprises, Directory /Non-Directory Enterprises and also some of the units coming under ASI. Value added per worker or per enterprise varies from one sub-sector to another. Hence, unless the sub-sectors are properly identified and a separate survey conducted for each, proper GDP estimates can not be obtained. It is, therefore, necessary to do this segregation so that newly emerging areas can be handled without any problem.

**New Areas of Study**  Due to modernisation and globalisation of economic activities, the existing activity scenario is fast changing and new areas are developing and contribution from them to GDP requires to be estimated. Some pilot studies may be undertaken to assess the effect of newly emerging areas. In order to assess the quality of statistics that are regularly collected, studies can be undertaken from alternative sources, if available, to see the difference if any and adjustment may be made in the data for making refinement. With the emergence of the new areas of economic activity that will contribute to GDP, the employment scenario will also get changed and this situation has created a high pressure on labour flexibility. In order to assess the employment situation a pilot study requires to be taken up to see the impact of labour flexibility in India.

**Instrument of Data Collection**  Schedules through which industrial data and socio-economic data are collected, are very bulky and require lot of time to be filled up and then processed. In order to ensure timeliness and quality of statistics, the schedule size should be reduced to one to 2 pages only with a proper format for recording data so that filling up of the schedule takes very little time; and collected data can be scanned or entered for on-line transmission. This will tremendously reduce processing time and results can be brought out within a month after completion of a survey. However, a type study with a small sample can be undertaken with the large-sized schedule to assess the quality of some estimates in case of doubt.

**Establishment of Computer Network in DoS**  In order to be self-sufficient in data processing, data communication and data dissemination, it is necessary that the DoS establishes a national computer network so that it need not depend on any outside agency and work can go on uninterruptedly. This

will help not only in the quick transmission of data on-line from various field offices to the data processing centres, but also in making on-line queries for data verifications. This will also help in working out a calendar of statistical events like planning a survey, release of compiled results, etc., and in monitoring the same.

**Training of Statistical Officers**  All personnel engaged in statistics are to be trained in the field before their posting. In general, all are to be trained in the use of computer so that in due course paperless statistical system can be introduced in collection, tabulation, and dissemination of data. Besides, all senior officers are to be given training once in three years, so that they can keep themselves abreast of recent developments in their own fields as well as related fields. For this, Training Divisions of the DoS should interact with Universities and also with applied statisticians engaged in other places, on a continuous basis, so that they can draw meaningful syllabi in various disciplines of statistics and related subjects and can identify the courses that will be suitable at various levels of officers. The proposed system of training will instil a sense of dignity and responsibility among the officers. As a result, a new dynamism in work-culture will develop and the statistical system in India will flourish.

# ON SOME ASPECTS OF STATISTICAL DATABASE
# IN POST-INDEPENDENCE INDIA

## Vaskar Saha[*]

## Introduction

The statistical database of the country had not been very good when India acquired independence from the British colonial rule in August 1947. At that time, the then government departments had been collecting data more or less regularly on certain essential topics like agriculture ( Land Records and Revenue Department ), population (Census Department), external trade ( Directorate General of Commercial Intelligence and Statistics), several geo-morphological information ( Survey of India ), etc. After acquiring independence, our first Prime Minister Jawaharlal Nehru felt difficulties in reconstructing the economy of independent India owing to acute problems of data and information gaps, particularly on certain macro-economic aspects like savings and capital formation, debt and investment, price, level of living, etc., and also on some other aspects of the people of India, in detail, like education, health, employment and poverty, etc. For the task of creating statistical database, Nehru then obtained suggestions and active assistance from Professor P.C.Mahalanobis in filling up quickly the data-gap through sampling. For this purpose Professor Mahalanobis proposed to the government the creation of a survey organisation to conduct regular multi-purpose sample surveys to meet the requirement of socio-economic planning and policy-making. The government of India agreed to this proposal and the National Sample Survey (NSS) was created in 1950. Subsequently the Central Statistical Organisation (CSO) came into being in 1951 for the purpose of coordination and standardisation of the statistical activities of the various statistical departments in vogue under the government of India. In order to fulfil its objectives, the CSO tried to establish and coordinate the activities of the statistical departments spread over different geographical areas in India. After the reorganisation of the States in 1956, we have the State Statistical Bureaus (SSBs) in almost all the State governments and Union Territories. Subsequently, the SSBs were renamed as Directorates of Economics and Statistics (DESs). With the passage of time various Central Departments and Ministries also expanded their activities, quite rapidly, in respect of collection, processing and publication of data and information needed for the formulation and review of various development programmes. With the tremendous increase in the data compilation activities over the last 45 years or so, some of the ministries and departments started facing problems in timely release of finally processed data. Over and above, there has not been much progress so far for the creation of an integrated database for quick dissemination of data to the concerned users. An attempt is, therefore, made in this paper to suggest a few points towards improving and reorganising the existing statistical database more effectively mainly from the angle of administrative control.

## Statistical Network for Dissemination of Data

With the recent spectacular development in computer technology in respect of both hardware and software, most of the data, collected either through complete enumeration or through sample surveys, are being processed with the help of computers. But at present there is no integrated database either at the national or at the State level, on the basis of which the concerned users in

---

[*] Director DPD & SDRD, NSSO, Calcutta. Views expressed in the article are those of the author and not necessarily of the organisation to which he belongs.

different organisations can have a ready access to the required data. Thus we can appreciate that there is an urgent need for the development of a menu-driven, user-friendly database arranged heirarchically from broad heads to individual items with the data already in existence. It is true that the chronological reference period of the database for the different aspects cannot be made similar as the availability as well as the periodicity of data release on different subjects are not the same. However, as the reference period of the data will be mentioned in the database, the concerned users will not face any difficulty. Introduction of such an integrated database will also ensure timely release and maintain reliability of data generated by various organisations, through their effective use by various applied scientists, research workers and other users who would be constructively critical on the available database, and thus help its improvement. Further, if the data on different subjects are available through the same database, a statistical network can be created through which the entire database could be made available to the users throughout the country with help of a computer network using dedicated P & T line or through satellite. As this computer network would be primarily dealing with statistical information, it may be named Statnet. A task force consisting of a few experts may be formed for the creation of the integrated database in a time-bound manner. The task force may also be entrusted with the responsibility to provide guidance for the maintenance of the database as well as for the methodology to be followed in incorporating the future survey results and the related data in the database. After the development of the database, possibility may be explored for establishing Video-Conference Facilities (VCF) through the computer network with the important Central and State departments and ministries. With the introduction of the VCF in the computer network, the database could be improved substantially as problems faced by any data processing agency during creation or maintenance of the database, could be solved easily through instantaneous discussions with the aid of VCF. Further, as the discussions could be held through this facility with its in-built advantages, the movement of scientific personnel could be reduced substantially, and ultimately the VCF should be more effective and economical.

## Coordination and Standardisation of Statistical Activities

Coordination and standardisation of statistical activities are the two major *rôles* played by the CSO. But the CSO has now very limited control on the unification of statistical activities conducted by other departments and ministries, leading to diverse estimates on the same topic. It is thus becoming increasingly difficult for the CSO to perform its said important *rôles* efficiently. In order to improve this situation, it is proposed that the Department of Statistics should be brought under the Cabinet Secretariat, shifting it from the Ministry of Planning and Programme Implementation. This step would enhance central control on other data sources originating in different departments and ministries on the same topic. Further, in order to have a better coordination among the various departments and ministries, we suggest that a very senior statistician of the CSO, for example, the officer in the administrative hierarchy just below the Director General, CSO, be given the full responsibility for the coordination of all statistical activities of this country. The statistical activities currently being undertaken by the various departments and ministries may, however, be continued. Here it is understood that a system of dual control - both by the CSO and by the concerned ministry or department should be necessary. To facilitate effective coordination we further suggest that there should be at least four joint meetings, consisting of all the concerned ministries and departments, in a year along with the CSO. The advantage of the proposed system is that it would be much easier for the CSO to ensure uniformity in concepts, definitions and coverage, reliability and timely release of data. Further, unnecessary duplication of work, if any, can also be avoided through the implementation of this system, granting, however, that the coordinating officer of the CSO would have the full responsibility for sanctioning any new major statistical activities to be undertaken in future by any department or ministry of the Central or State government.

## Advance Calendar for Release of Data

If the data on any topic cannot be released within a reasonable time after the reference period of survey or compilation, the compiled information may become outdated, greatly reducing its importance and relevance. But currently some of the data compilation agencies with their present system of processing that is devoid of appropriate utilisations of advanced computer technology are not in a position to finalise the data in time. In order to improve this situation, it is proposed that the calendar for release of data in respect of at least important statistical activities be released compulsorily through public notification at the beginning of each calendar year, so that processing personnel may have the time frame in advance on the responsibility to be accomplished. The infrastuctural facilities required for the collection and processing of data must, therefore, be provided to the concerned agencies. Further, a close fortnightly or monthly monitoring system should be introduced to achieve the target date of release fixed at the beginning of the year. This aspect of timely release of data has now become all the more important, since India has recently decided to participate in the Special Data Dissemination Standards (SDDS) established by the International Monetary Fund (IMF). According to the SDDS, it is obligatory on the part of the participating countries to release data regularly within the stipulated time period as decided by the IMF on certain important aspects to be identified by the participating countries from a selected set of topics fixed by the IMF. Further, any participating country cannot violate the time frame of data release without having shown sufficient reasons for such violations on any particular topic.

## Comprehensive Statistical Act

Statistical data and information as compiled from the fields are at times incomplete, because of non-cooperation from the respondents in the field. In order to avoid this difficulty, there should be an appropriate kind of Statistical Act, so that respondents must compulsorily cooperate in divulging information to the authorised statistical investigators. But the Statistical Act as prevalent at the moment is applied only to some limited areas of compilations. And such Acts are not even powerful enough for timely capture of data. Thus there is an urgent need to review the existing Acts for the collection of statistics in order to bring out some sort of comprehensiveness in the Statistical Act that could provide sufficient administrative power to the statistical investigating authority for collecting data from an individual or a household or an enterprise or any organisation in the field. Suitable penal action for non-cooperation, intentional misinformation or refusal to provide information by the respondents to the statistical investigating authority may also be incorporated in the Act. However, it is to be ensured through the Act itself that the information furnished by the informants will be kept strictly confidential by the statistical authority.

## Methodological Studies for Collecting Data on Unorganised Sector

In our national economy, a very large part of human labour is engaged in the unorganised sector. As such, this sector plays a very important *rôle* in terms of its contribution to the gross domestic product. However, activities generated in this sector have expectedly a high mortality rate. Again, the concerned establishments or enterprises on such unorganised activities are sometimes floating and do not normally maintain any book of accounts. It is thus extremely difficult to collect reliable information from this sector. Thus there is an urgent need to conduct in-depth methodological studies for collecting reasonably dependable information on this very important sector of our economy. While conducting the methodological studies, the knowledge of experienced persons engaged already in these fields and also the previous survey results, if any, should be fruitfully utilised. For example, we know from our earlier experience that if we make an attempt to

collect data on both income and expenditure of any establishment in this sector, the information on income is generally understated while that on expenditure is overstated. Again, from our experience we notice that sales data are understated very often in the unorganised trade sector. The present practice is to collect both purchase and sales data directly from the same entrepreneur in this sector, with a view to estimating the value added at the national or sub-national level. Instead of this practice, we can think of avoiding the direct collection of sales data and ask for the data on purchase and trade-margin directly and derive the sales value from these. It should be noted that the estimates of value added as derived from the understated sales data are often negative. This is impossible in the long run, since no entrepreneur will be in the business incurring continued losses. However, a detailed pilot study is needed to test this methodology for collection of proposed data in the unorganised trade sector. Such methodological studies should also be extended on other unorganised activities (such as transport, construction, service, etc.) for arriving at appropriate items for data collection.

## Benchmark Study for the Annual Survey of Industries

The Annual Survey of Industries (ASI) is conducted every year to collect information on certain important characteristics of the industrial units registered under the Indian Factories Act, 1948. Under this Act, all the industrial units employing 10 or more workers with the aid of power or 20 or more workers without the aid of power are to be registered compulsorily. In the ASI, all the registered industrial units are surveyed not exactly annually, but within a time span of three years. This is primarily because the number of industrial units is quite high and it has not been possible to cover all units every year on account of the limited resources available for this purpose. Again, while designing for the collection of data for this sector, the industrial units are divided into two groups, namely, census and non-census. All the industrial units employing 100 or more workers are covered under the census sector. In case of industrially less developed States and Union Territories, however, all the industrial units are also surveyed every year irrespective of the employment size of the units. Further, if the number of industrial units is less than 20 at the 3-digit level of industrial classification in a State or Union Territory, the entire 3-digit group is covered in the census sector. All the remaining units are covered in the non-census sector. The industrial units belonging to the census sector are covered every year. The units belonging to the non-census sector, on the other hand, are covered on a sampling basis. However, the industrial units under the non-census scheme are all surveyed over a period of three years, as mentioned earlier. Thus the statistical features of the registered units can only be estimated for a particular reference year, but the actual features can not be known exactly for any year. In order to know the exact statistical features of the registered industrial units at least once in a time span of, say, five years, it is desirable that all the industries available in the list of factories are surveyed, say, once in every five years as part of a benchmark survey. If complete enumeration of registered industrial units is introduced at five-year intervals, there should be some sort of simultaneous sample surveys of registered industrial units on a monthly basis in every year. If the benchmark survey is accepted for the ASI sector, the list of factories could be suitably updated on the basis of the benchmark survey results and the survey design of the ASI may accordingly be modified, if necessary. With the introduction of the system for generating monthly data on the ASI sector, we should ensure that the monthly survey results must be released within a month after completion of the survey.

## Divergence in Similar Data Released by Two Agencies

When two government agencies are involved in the collection and processing of data on similar subjects, appropriate actions need to be taken by these two agencies to maintain uniformity in the coverage, concepts and definitions as well as the methodology followed for conducting the

surveys. Even if the same definition can not be followed by an agency to meet its special requirements, appropriate provisions need to be made in the survey design to generate comparable estimates on a common characteristic or item by both the agencies. This is required to minimize the divergence in the estimated values of similar characteristics between the two sources of data on the same item. It is also necessary that the methodology, data source, coverage, concepts, etc., and particularly the limitations of data, if any, should be clearly brought out for the benefit of the concerned data users.

## Enhancement of Computer and Training Facilities

During the last few years, there has been tremendous development in the computer technology. In order to avail of the full benefit of the recent improvements in both hardware and software technologies, there is a need for enhancing the computer hardware facilities in both Central and State government offices. However, while taking decisions on the operating system and the relevant software packages, the needed compatibility in respect of both data processing and interchange of data must be assured among the concerned offices. Further, E-mail and Internet facilities should be provided at least to the heads of all the concerned statistical offices. While planning for the enhancement of computer facilities, appropriate training programmes for the computer personnel need to be formulated. Most of the training programmes that are now in vogue provide only a generalised exposure on certain software, that may not find full utilisation in the day-to-day functioning of the Central or State government offices. But instead of such a general exposure, it is desirable that the requirements of the participating offices should be assessed first by the training institute and then job-specific course content should be chalked out and imparted to the participants, so that, at the completion of the training, the participants should be able to perform their day-to-day work independently and efficiently.

## Development of Expertise in Some Important Fields

The NSSO has been conducting socio-economic surveys on different subjects since 1950. The subject coverage (as now in vogue) having been different from one round to another, the main activities like designing of schedules of enquiry, drafting of instructions manual, preparation of tabulation plans, etc., are subject to change over the years. To keep track of these changes, experienced statistical personnel are required to accomplish all the necessary tasks. Since the subjects coverage is different in different rounds, the officers having gained invaluable experience in the survey design and data processing over the years are expected to make significant contributions to this process as compared with a new officer in this field. Thus, there is a need for retention of the experienced officers in these offices so that their experiences could be fruitfully utilised for the maintenance of continuing tasks. It is, therefore, proposed that ISS officers who have served for about 15 years in the statistical system be allowed to specialise in the area where they feel themselves most capable of making significant contributions on the basis of their experience.

## Concluding Remarks

Certain important aspects of the improvement of database have been discussed above, mainly from the angle of administration and control of statistical systems. Technical details on these aspects, however, have not been discussed here for the sake of brevity. Again, the priorities on different aspects must be decided for the sake of resource limitations. But we have not pointed out such priorities on different aspects. Before implementation of the proposed line of improvement, priorities and technical details will have to be worked out. However, the proposals made above may have enlarged our vision on the possible improvement of the statistical database in the future.

# FIFTY YEARS OF THE NSSO - AN APPRAISAL

## Pranab Kumar Sen[*]

The National Sample Survey Organization (NSSO) had its genesis as a vital part, viz. the NSS wing, of the Indian Statistical Institute (ISI). It has emerged as an essential centre for the training of diverse statistical personnel and provides a model framework for the collection, maintenance and improvement of quality (and coverage) of statistical data (and information) systems pertaining to the complex Indian public as well as private sectors. The NSS wing had a profound impact on the Central Statistical Organization (CSO), at the Union Government level, and all the State Statistical Bureaus at the regional level, the office of the Directorate General of Commercial Intelligence and Statistics (DGCIS), and a bunch of other agencies of the Central and at the State Governments. It laid down the foundation of the much needed Indian Statistical Service (ISS). Even the (Indian) Five Year Plans (in the formulation of which Professor P.C. Mahalanobis played a major rôle ) could not have been successfully implemented without the able assistance of the NSS wing (NSSO). In this way, the NSSO has become a model for the Third World countries (on how to create an adequate and operable statistical database consistent with the available national resources as well as manpower). Indeed, the work of the NSS wing was an achievement of the foresight and enormous vision of its founder, the late Professor Prasanta Chandra Mahalanobis. In spite of other acute problems (such as achieving self-sufficiency in food and other necessities of life) India encountered at the time of attaining independence (from British rule), the NSS wing made tremendous progress at the very young stage; this was largely possible due to the intensive cooperation and collaboration the NSS wing had with the other units of the ISI. (Even the Research & Training School of the ISI contributed richly to the development of needed sample survey methodology that formed the operational platforms for the NSS wing.) In this broad sense, indeed, with respect to its constitution, professional as well as academic imputations, and operational plans, the NSSO is rather unique across the occidental as well as oriental countries; its accomplishments have been highly commended by parallel organizations not only in other developing countries but also in the industrialized nations.

In a global perspective, the creation and operational plans of the Food and Agriculture Organization (FAO) and some other divisions of the United Nations Organization (UNO) as well as the developments at the International Labour Organization (ILO) have also been considerably influenced by the mastermind of the NSSO. A significant number of highly skilled Indian statisticians having their prior training in India joined in various capacities these UN centers, and contributed richly in the development of statistical systems there; special mention may be made of the late P. V. Sukhatme, who reshaped the statistics division of the FAO completely. We are indeed very grateful to the outstanding researchers in the area of sample survey methodology who gathered in the two camps: Mahalanobis-Lahiri-Murthy in the ISI, and Panse-Sukhatme at the statistics division of the Indian Agricultural Research Institute. Not only has the NSSO benefited tremendously from their fruitful, application oriented statistical research, but also has the entire community of agricultural (and livestock) research scientists who have used their findings in actual field-work. Indeed, the separation of the NSSO from the ISI has the potential danger that in the event of less than adequate methodological research support, the mode of operability as well as capability might have to be compromised. Of course, the NSSO may have, by now, an adequate number of research workers to guard against such probable erosion.

In the United States of America, statistics is deeply rooted in all areas of applied research: health, economics, social as well as agricultural sciences. The Bureau of Labour Statistics, the Food and Drug Administration (FDA), the Center for Health Statistics, and a bunch of other organizations

* Professor of Statistics and Biostatistics, University of North Carolina, Chapel Hill, USA

(or agencies) carry out a complex of statistical activities, in more areas, than the NSSO does, mostly alone, in India. In Canada, Statistics Canada has a similarity of set-up with the NSSO. Though, the developments in Canada were primarily influenced by the US and the British Commonwealth set-ups, it would not be wrong to mention that the NSSO has had a profound impact on the current operations of Statistics Canada too. Many of the leading statisticians in that organization are of Indian origin and their training has greatly influenced the system. The basic difference between USA - Canada and India is that bringing the diverse set-ups in a diverse country like India, at a time when there were paramount food, health and economic problems, was indeed a difficult task, and the NSSO should be commended on the job it has accomplished during its first 50 years.

Let us look into the immediate past and make projections for the years ahead. India has achieved some sort of self-sufficiency in respect of its food and agricultural needs. However, in many other sectors there is a definite need to concentrate more. Among these, health (management and care), education, welfare, and environmental threats are to be taken up with utmost seriousness.

Due to increased domestic and industrial needs, the energy crisis has been labeled as a primary concern. A mounting consumption of gasoline and diesel products (for which India is far from being self-sufficient, and a greater part has to be imported from our Arab neighbours at a comparatively high cost) has created a spiral of economic insecurities on the one hand and on the other, has led to a battery of tremendous environmental and ecological imbalance problems. While our industrialization progress has contributed towards the "greening effect" (thinning of the ozone layer) and "environmental smoking effects", our big city-centers are so much clogged with automobile exhausts and environmental toxicants that the pollution level is dangerously high. There is no relief in sight! The Himalayan foot-hills, once ranked among the best health resorts, have lost their superiority to a great extent. While this has been partly due to the population explosion (and related problems), it may also have been due to our changing needs. Elimination of mid-to-high altitude vegetation level (for fuel consumption, dwelling and agricultural land) is not only causing a shift in the climatic pattern but also contributing largely to soil erosion due to landslides, etc., that occur now with a greater frequency than before. The operation of our oil-refineries has added layers of sulphur-toxic material to some of our historical monuments (to the Taj Mahal, for example) and also is increasing the health hazards at an alarming rate. Can India afford to have an energy plan without proper feedback from its national resources as well as its needs? Moreover, this has to be assessed not only for the immediate future but also for decades ahead! Who would be entrusted with this marathon job?

In USA-Canada, environmental problems are being encountered with greater awareness and sensitivity. Environment Canada and the Environmental Protection Agency (EPA) in USA are working together to handle environmental problems on a continental basis in North America. Statisticians are well involved in this enormous task. Environmental (E) mapping of various potential places (where environmental pollution is likely to be much higher) has been an important (initial) step toward a solution. Development of proper technology to minimize some of these toxicity-effects has also been under progress. In the health sector, cancer and heart diseases have been identified as major factors. Even the AIDS (HIV) has caused a global concern and much greater medical attention is needed to combat this unprecedented epidemic disaster. India, is by no means, free of these dangers and hazards. The World Health Organization (WHO) has placed India in the highest incidence-risk group with respect to the AIDS virus, and the projections for India are rather grim!

In India there are, by now, some central health institutes (such as the All India Institute of Medical Sciences, Institute of Immunology, National Institute of Health and Family Planning, mostly all in the Nation's capital), and also some others devoted to environment and ecology. But such research institutes are likely to have a greater emphasis on medical and health sciences, with persistent epidemiological as well as demographic components. Thus, they may be more inclined to deal with the epidemiological aspects, such as the incidence rate, mortality rate, etc., but not so

much with how such facts and figures have to be compiled for the nation as well as its various sectors, characterized by demographic and socio-economic diversities. Therefore, there is a need to develop a system of data collection and information gathering that goes far beyond the vital statistics records that the office of the Registrar General of India compiles on a routine and periodic basis. This is definitely a more complex, more delicate task that needs close supervision by trained and skilled statistical personnel, who could come up with adequate statistical measures and concepts to assess the dimensions of the problems, develop suitable statistical analyses to examine the risk-factors in a proper fashion, and improve the data-collection and monitoring systems. It is our sincere hope that the NSSO, with its acquired expertise in the area of agricultural and socio-economic information retrieval, would be able to formulate the necessary steps to ensure that this essential national task is properly taken up, and accomplished with satisfaction of all concerned. Such an accomplishment may depend not only on the collaboration of scientific workers in the field of health and environment with statisticians, but also on the cooperation of all the allied agencies so that a general consensus can be reached and a project manual be framed with safeguards from all sides. This would probably call for adding new divisions or sectors to the NSSO framework, and would have to address the health and environment problems more adequately. Another important aspect of this emerging field of applied statistics is public awareness. In western countries, the discipline of public health has come out as a broad avenue of examining various aspects that range from health administration and health policy, health education, disease prevention and health promotion, maternal and child health, nutrition, epidemiology and environmental sciences to environmental health sciences and a variety of allied sub-disciplines. Biostatistics has emerged as a binding force, and it has been recognized as an essential ingredient in the pursuit of many of the above fields of study and practice. Being in this position, Biostatistics has been endowed with the responsibility of data collection, planning of public health studies, data-monitoring and quality control, statistical modelling as well as statistical analysis of acquired data sets. The conventional sample survey methodology that pertains to agricultural and socio-economic fields may not be properly applicable in this field of bioenvironment and public health. Sampling designs are quite different and, since human beings are chosen as sampling units, medical ethics and other considerations may generally impose additional constraints that complicate the plans and invalidates the use of conventional sampling designs. A variety of novel methodology has emerged in the recent past to address some of these issues in an adequate manner. In India, there are a few public health institutes (the oldest being in the heart of Calcutta). However, I have no idea how much they are involved in the statistical aspects of data collection, quality-inspection and analysis, and drawing of statistical conclusions. Any method that blindly incorporates a very simplistic model is likely to be non-robust and improper. Therefore, there is a need to make proper assessment of the basic regulatory set-up before implementing statistical protocols to draw conclusions. Should the NSSO march into the public health sector too and enhance its adaptability?

In two years, we would be marching onto a new millennium, and public health and environment have been identified as the two most basic concerns for the mankind, not just in the Indian subcontinent. As such, the NSSO having its rich expertise and heritage should not hesitate to undertake such added responsibilities, and enlarge its sphere of activities for greater benefit of the country. This is a challenging task for statisticians, and the NSSO should provide the leadership in evolving proper sample survey methodology that is very much needed in this country and at this point of time.

# IMPROVING THE STATISTICAL DATABASE IN INDIA
## - SOME SUGGESTIONS

D. Sinha[*]

The importance of statistics in formulating development plans had been recognised right from the time India achieved her Independence. In January 1949, Prof. P.C. Mahalanobis was appointed Honorary Statistical Adviser to the Cabinet and a few years later concurrently worked as a member of the Planning Commission. In 1954, the Planning and Statistics Division of the CSO started functioning as the Statistics and Survey Division of the Planning Commission to provide the desired statistical support in the actual planning process. It was through this Division of the Planning Commission that the CSO coordinated the statistical programmes drawn up by the Central ministries and State governments. Since then, this organisation, under the Department of Statistics, has been holding an apex position in the official statistical system of the country. Its current responsibilities have increased manifold and include among others (i) coordination of all statistical activities at the Centre and in the States with a view to identifying data gaps or duplication of statistical work, (ii) development and maintenance of norms and standards *vis-à-vis* concepts and definitions and methodology of data collection as well as their presentation, (iii) liaisoning with UN and other international organisations on statistical matters which include furnishing of data to those organisations, (iv) preparation of National Accounts Statistics, (v) Conduct of Annual Survey of Industries, Economic Census and Follow-up Surveys and (vi) training of statistical personnel. The National Sample Survey Organisation through its large-scale Socio-Economic Surveys on a wide variety of topics such as household consumer expenditure, employment-unemployment, literacy and education, health care, public distribution, housing conditions, land holdings, debt and investment, etc., provides valuable input for national planning and policy formulation.

## 1. The Problem

It is true that, as a result of improvement and expansion in the statistical activities during the last few decades, a larger volume of data is now available in almost all important fields. With proper functioning of the current system, it should have been possible for the Department to devise an efficient procedure by which data could have been made available to the users within a reasonable time frame. But unfortunately, the statistical system in India is yet to fully achieve this objective.

A number of committees, commissions, conferences and seminars, wherein official and non-official agencies, research workers and other experts participated, have identified gaps in the database of the Indian economy indicating deficiencies in the statistics being produced. While corrective action on some of the deficiencies have been taken by the DoS, a few of the issues still remain unresolved. This paper proposes to focus on some of the critical issues that can significantly contribute to improving the statistical database of the country.

## 2. Time Lag

Timely availability of statistics is the essence of efficiency of the National Statistical System. It is crucial for planning and decision-making as well as for effective management of economic affairs in government departments. Statistics with a long time lag may not be of much use to the policy-makers. If the statistical system is to function effectively, ways and means have to be devised to reduce the time lag and provide the policy-makers with some basic useful statistical material to facilitate their decision-making.

[*] Joint Director, DPD, NSSO, Calcutta. Views expressed are those of the author and not necessarily of the organisation to which he belongs.

## 3. Simplification of Schedules

The schedules/questionnaires used for collecting statistical information should, as far as possible, be short and simple. It needs to be ensured that data on items which are not included in the tabulation programme are not collected. This would not only help to reduce delays in compilation and processing but also would minimise informant's fatigue and resistance to supplying such information.

## 4. Computerisation

Although an increasing volume of statistical information is being collected both by the Centre and the States, the existing arrangement of data processing in the States is not adequate to make the results available with the desired speed and accuracy. A recent study by the Institute of Applied Manpower Rearch(IAMR) based on the analysis of information of four Socio-Economic rounds - Nos. 48 to 51 of the NSS - received from 26 States revealed that between 58 % to 65 % of the States do not at all process the data collected on a matching basis, 69 % to 89 % do not generate any tables and about 50 % do not have the EDP facilities. This is resulting in a huge loss of information collected by deployment of a substantial part of the Central grant to the participating States. The main idea behind the States' participation in the NSS programme was to increase the sample size to provide more disaggregated results at regional level within each State by pooling the results of the Central and State samples. With the infrastructural facilities of the States, as they are today, the pooling of results has neither been achieved nor is there any possibility of such an achievement in the near future. It may, therefore, have to be considered if such a scheme should at all be continued or the processing facilities of the NSSO, DPD, currently entrusted with the tabulation of the Central NSS data, have to be extended to those States which are significantly lagging behind in this respect.

## 5. Provisional Estimates and Summary Tables

In addition to preparing the full complement of statistical information on a comprehensive basis, all statistical agencies should devise a system to provide a series of quick provisional estimates and summary tables. In addition, suitable indicators can be developed and made available in each sector for periodical monitoring and policy adjustment. This would help to meet the needs of a large number of data users including policy-makers whose primary interest may not be in detailed results but only the main results in a summary form. For important surveys, where processing of the entire data collected causes delay in release of results, advance tables based on a sub-sample of collected data can be generated, wherever feasible, for the use of planners and policy-makers.

## 6. Creation of a Data Dissemination Centre and a Central Database in the CSO

The Department of Statistics in the recent past has been giving wide publicity to the various statistical series being maintained by them. To provide easy access to the administrators, economists and various other users of the data available at the national level, a comprehensive central database can be set up at the CSO in respect of data available in the CSO, NSSO and the Central ministries. This can be extended to the States in due course. The prepared database should keep detailed information on availability of different types of data, periodicity, sources along with some explanatory notes on concepts, coverage, time-lag, level of aggregation/disaggregation and so on. In order to keep this central database up-to-date, data released for official use should be immediately incorporated into the database. This centre, for better dissemination capabilities with various external agencies should not only be connected with standard networks like the INTERNET or NICNET but should also be equipped with adequate reprographic facilities to make selected data available to officials and other data users according to their needs.

This database can subsequently be extended to cover various administrative data required by the Central ministries and other agencies of the Central government from time to time. It may not be out of place to mention that the Fifth Pay Commission in a note on the "inadequacy of data required for pay commission work" has referred to the time lag in collection of administrative data and also to the contrary trends in size of the Central Government employees being collected by the different agencies of the Central Government like the Department of Personnel and Training, the Directorate General of Employment and Training, Department of Expenditure and Department of Economic Affairs. They have also referred to the incomplete information available in respect of certain other items like pay and allowances of the Central government employees, pension and retirement benefits, etc.

The CSO being the nodal agency and entrusted with the responsibility of coordinating the statistical activities of the various ministries/departments, the work of such compilation should ideally be taken up by the CSO under the DoS. However, since the information is to be compiled on the basis of data supplied by the numerous offices of the Central government located across the country, the *modus operandi* and the periodicity with which the data should be collected have to be decided keeping in view the requirements and the complexities involved in this operation. The DoS may also have to devise a suitable coordination mechanism by which any delay/non-receipt of information from any ministry /department can be immediately attended to. This additional activity can be taken up by the CSO by transferring a few of the posts already sanctioned in the various ministries/departments and who are engaged in compilation of this information.

## 7. Discrepancies in Data Series

Discrepancies between statistics relating to the same subject released by different agencies are quite common. These discrepancies arise mainly on account of (a) a lack of uniformity in concepts, definitions and methodologies followed in data collection by different agencies, (b) differences in coverage and (c) differences in the length of reference period and/or period of survey. This problem could be solved to a great extent with improvement in the mechanism for coordination of statistical activities. As far as possible and as a matter of policy only one set of data on a particular item should be available and a single agency should be authorised to release the same. However, in a situation where there is a multiplicity of agencies, all involved in the collection and release of data on the same subject, a reconciliation statement should invariably be prepared and published along with the main data series.

## 8. Reliability

The quality of primary data is sometimes so poor that inferences based on them become unreliable and misleading for any decision-making process. It may generally be seen that publications containing statistical data do not adequately indicate the limitations of different data series. In many cases, the reports containing the results of sample surveys do not provide standard errors of the estimates. In the absence of error estimates, the data series and also the results of the sample surveys lose their credibility. It needs to be ensured that every statistical series released clearly indicates the limitations, if any, and further all the results based on sample surveys should be accompanied by standard errors of the estimates.

## 9. Augmentation of Statistical Personnel

The statistical offices need strengthening with regard to both manpower and expertise in specialised field so as to possess dynamic leadership and the desired level of analytical capabilities. It may be seen that in some of the ministries, different departments have independent statistical units whose work at present is not well coordinated. Since statistics is a vital input in the decision-making process, coordination of statistical activities within a ministry is considered very important. It is therefore necessary for each major ministry where integration of statistics of different subjects

handled by the ministry is essential for taking policy decisions, to have a whole-time statistical adviser at a sufficiently high level who can directly report to the seniormost functionary of that ministry. This will not only ensure that the entire statistical work of the ministry is well coordinated but will also provide the ministry the right type of statistical advice needed for formulation and planning purposes. Incumbents to the posts of statistical advisers should be professional statisticians with sufficient experience in the field of official statistics.

## 10. National Advisory Board on Statistics

To make it an effective catalyst in bringing about qualitative changes in the statistical database of the country, the National Advisory Board on Statistics has recently been reconstituted by the DoS. The NABS is expected to concentrate on the larger issues of developing a strategy for ensuring the best possible functioning of the entire national system.

In order to ensure adoption of standard concepts, definitions and methodology and avoid duplication in data collection, it should be made obligatory on the part of statistical agencies to obtain clearance from the NABS before launching any large-scale statistical operation funded by the government. However, since it may not be possible for the NABS to screen each and every proposal, only those exceeding a prescribed cut-off monetary limit may be considered.

## 11. Subject Panels

A systematic and sustained effort needs to be made to bring about uniformity in concepts, definitions and methodologies used in all data collection programmes. In a dynamic situation, the concepts and definitions need to be reviewed from time to time to take into account the requirements of the data users while ensuring that comparability is maintained as far as possible. The CSO may consider the setting up of subject panels for all important areas on which data are concurrently being collected or are proposed to be initiated either at the Central or at the State level. The subject panels, apart from having representatives from the CSO/DoS, should have subject-matter specialists from administrative agencies, universities, research institutions and user organisations.

The subject panels, will be responsible for laying down of standards in different subjects and ensuring their adoption and use during collection of data, assessing the statistical requirements of the concerned subjects, advising on matters relating to the collection, tabulation and dissemination of data, evaluating the problems and suggesting remedial measures for the improvement of quality, reliability of data and timeliness of release of results. The NABS can oversee the working of these subject panels.

## 12. Conclusion

The issues mentioned above are a few areas of concern which deserve serious consideration. With the opening up of the Indian economy, there has been a perceptible change in the requirement of data users. It is felt that the above measures will help create a sound statistical database at the Centre and help the CSO meet the challenges of the future.

# EASY ACCESS TO SOCIO-ECONOMIC DATA

## T.N. Srinivasan[*]

"India stands out from other countries in terms of its tradition of data collection and its pioneering of many of the techniques of data analysis which have now become common currency throughout the world...**Yet access to existing data resources remains highly restricted**. For example, few researchers in India have received permission to analyze NSSO data at the unit-record level. There are an enormous number of urgent questions which could be investigated if such data were accessible, and there is certainly no shortage of well-qualified researchers in India who could undertake such studies." [World Bank, *India: Achievements and Challenges in Reducing Poverty*, Report 16483-IN, May 27, 1997 (emphasis in the original)[1] .]

I would like to take this opportunity to discuss briefly about an issue that has always been a central concern of all the researchers, namely, socio-economic data in their many dimensions, in particular timely availability, accuracy, adequacy and, above all, access to researchers. Many of us know that the Indian Econometric Society held its first conference on the database of the Indian economy as far back as 1972. Subsequently six more conferences were held, the last one more than a decade ago. Each conference was devoted to one particular area, and several papers analyzing the data quality and availability in that area and making recommendations for improvement were presented. The Society has published five volumes of conference papers. This series of conferences was unparalleled, not only among developing countries, but also in the entire world. Thanks to former Prime Minister Rao and former Finance Ministers Dr. Manmohan Singh and P. Chidambaram, the Indian economy is being integrated with the World economy, after nearly four and a half decades of disastrous insulation from it. With this sea change, the data needs of the economy have also changed. At this crucial juncture I do think it would be appropriate to examine whether the Indian Statistical System is using scarce resources in routinely collecting and compiling data that serve simply as inputs for international data compilations, based on frameworks that might be less relevant for the Indian economic reality. To take just one example, I would argue that many elements of the framework of the United Nations System of National Accounts have little to recommend themselves for the purposes of describing or analyzing the economic structure of India or other similar developing economies. Resources saved by not compiling less useful data could obviously be better utilized in collecting more appropriate and needed data.

I have argued elsewhere (Srinivasan 1994) that the United Nations, the World Bank, the FAO and other such organizations, instead of publishing inadequately documented data that they do not themselves collect and which are spatially and temporally non-comparable, and conducting expensive surveys that apparently leave no lasting impacts on survey countries in terms of survey capability, should instead spend the resources involved in creating data gathering and analytical capabilities in developing countries if they do not exist, and strengthening them where they do. The unsatisfactory situation with respect to socio-economic data led me to organise a conference in 1992 at Yale University on Database for Development Analysis with the financial support of the Ford Foundation. The papers presented at that conference were published in a symposium issue of the *Journal of Development Economics* in June 1994. The papers describe in some detail the problems with internationally and nationally compiled data. The May 1996 issue of the *American Economic*

*Review* also contains four papers on the reliability, comparability and adequacy of international macro and micro economic data.

Many of us are aware that based on the panel data collected by ICRISAT in Hyderabad for a period of less than a decade, covering about 10 villages of Andhra Pradesh and Maharashtra, and from less than a thousand households in all, a large number of papers using sophisticated econometric models of household behaviour have been written, largely by American academics, and published in major journals of our profession. There is no doubt that these papers have advanced our knowledge of the behaviour of poor rural households. Yet as an Indian I feel outraged that the far more extensive household data collected by the NSS have not been so analyzed. The reason for this is simple: the cumbersome and restrictive procedures that govern the access to unpublished household level data. The current rules of access are as follows:

(a)    Permission to access data relating to small areas or few districts can be accorded to individual workers provided that:

i) the user would have to give a written undertaking that he/she would obtain prior approval of the council before publishing any results from NSS data supplied to the user;
ii) the paper prepared would be offered first for publication in *Sarvekshana*, the official journal of the NSSO; and
iii) while publishing the results based on NSS data, the user would acknowledge the source of data.

(b) Access to NSS data on a large scale would also be permitted in specific cases provided the research workers are prepared to work under the aegis of the NSSO and with the above-stipulated conditions.

Those who devised these rules honestly believed not only that NSS data are secret, but also that they are duty-bound to ensure that NSS data were not misused. It is worth remembering that the tax payers of this nation finance the NSSO and that the thousands of poor households who provided data over the last four decades did so willingly and without compensation. They did not insist that the information they supplied be kept confidential - on the contrary, such a thought would never have occurred to them. If any group had any right over the NSS data, it is these poor households. It is not at all proper for the authorities to have restricted access to the data. Of course, it is not my argument that access should be free of cost but only that it be free of restriction on use.

So much for the past. Fortunately for all of us, the Department of Statistics itself has recognized the problem with data access and the need to make statistical activity user-oriented. Early in 1996 it held a meeting of prominent academics and researchers to discuss the issue. The agenda paper for the meeting candidly admitted,

> "There has been no official policy for promoting easy access to unpublished data for users. Consequently, different Central Government Departments have reacted differently to the data needs of various users. All of them distinctly exhibit services inhibition in parting with unpublished data, in spite of the knowledge that the reports produced by them generally use only a small part of the large data generated by the statistical system.
>
> The above-stated inhibition in dissemination of unpublished data exists notwithstanding the fact that the huge national resources in collecting (unprocessed) data have been invested."

The paper went on to point out that the restriction that unpublished data cannot be given to any individual including interior user, resident abroad or any international user outside the country is

"devoid of rationality and directly contradicts the policies of openness [and] are the conditions that any paper written using NSSO data is to be first offered for publication in *Sarvekshana* inhibits free and fair use of unpublished data". After this welcome, though very belated, recognition of the ridiculousness of the existing restrictions, it then added that these restrictions had been violated: some published their papers based on clandestinely obtained NSSO data, and others who obtained them legally from NSSO nonetheless published their papers elsewhere without giving *Sarvekshana* the first right of refusal. I am not sure what the authors of these papers intended by drawing attention to these violations - whether they meant to reinforce their criticism of restrictions by showing they invite violations or whether they wanted to include safeguards against violations once some new rules are established. I hope it is only the former.

I suggest that the simplest and fairest rule of access to data is to make them available at cost to any and all. For example, the World Bank sells diskettes of socio-economic data most of which were supplied to the Bank by the member countries in the first place. There is no reason why, with the technology available today, it should not be possible to make individual household level data for all the 100,000 or so households canvassed in our quinquennial surveys available in a compact and reasonably-priced way. The precise identity of each household can easily be suppressed to maintain anonymity and confidentiality.

It has been argued in the past that some relatively less well-endowed research institutions in India and individual scholars of modest means may not be able to afford the fee, if one is charged for access. This is a specious argument - adequate resources can be provided through the ICSSR or other agencies to provide grants for worthwhile research to institutions and scholars so that they can acquire the data from such grants. Even more specious is the argument that foreign scholars with better funding and research facilities would forge ahead of less well placed Indian scholars in completing their research with NSS data and publishing them. Again, this should be an argument for upgrading the research facilities available to Indian scholars rather than for preventing foreign scholars from access to NSS data. I should add that these arguments, which could be generically called "level playing field arguments" are almost always specious whether they are advanced by the United States Trade Representative against Japan or by some Indian scholars or some research institutes against others at home or abroad. Another specious argument is that users may not understand the sample design and other aspects of NSS data and misuse them. After all, each scholar who publishes a paper is responsible for the analysis contained in it and risks his or her scholarly reputation on it. Whether or not the data have been correctly used and interpreted should be left to the peer review process of the profession at large and not entrusted to the NSSO bureaucracy.

**Reference**

Srinivasan, T.N. (1994), "Database for Development Analysis: an Overview", *Journal of Development Economics Vol. 44, No.1*, June, pp, 3-27.

*Editors' note*: Data dissemination procedure for socio-economic surveys has been changed since 1997. Presently the detailed data as well as reports prepared by NSSO are being provided to the users at a reasonable price. The compulsion to submit for publication of any paper using NSS data first to *Sarvekshana* has also been waived.

# FUTURE SHAPE OF STATISTICS

## S.G. Tiwari[*]

## 1. Introduction

Statistics should serve as barometers of the performance of the economy over time and space. They should be based on international standards adapted to meet the national requirements and should be adequate to meet short-term, medium-term and long-term policy and planning needs. They should cover all aspects of the environment as well as social, sociological and political life. Statistics developed should also be reliable. The reliability is to be judged from the perception of the people and the society.

In realising these objectives, work has been in progress in the country at the Central, State, regional and local levels for the past several decades. At the Central level, the work has been undertaken by the Department of Statistics in collaboration with various ministries of the Central government, national statistical institutions, universities and individual research workers.

There has been tremendous progress in different fields. However, there are several areas in which further work needs to be initiated. Work in several fields already covered needs to be reviewed with a view to improvement. Institutional arrangements also need to be reviewed for their adequacy to meet the current and future needs.

## 2. Statistical Work

Among others, the current work on statistical development at the centre is being done by two main organisations, namely (1) the Central Statistical Organisation (CSO) and (2) the National Sample Survey Organisation (NSSO). Over the years they have expanded their activities so as to cover varied fields.

## 3. The CSO

The CSO has developed statistical standards. It has undertaken coordination and promotion of statistical activities of various ministries of the Central and State governments. Its functions include maintenance of statistical standards and liaison work with international organizations. Liaison work with international statistical agencies requires not only that the CSO should adapt their standards constantly, but also that it should undertake experiments to make them suitable for the country. For their applicability, they need to be modified to suit our requirements.

The CSO should act as the fountain-head in the field of statistics. It should command respect due to its expertise, which should impel others to seek advice of the CSO. The staff of the CSO mostly consists of the personnel recruited to the Indian Statistical Service. It is good that they move from one job to another and are exposed to different disciplines. However, there should be a core staff in each discipline which should also be engaged in research. They should be retained and promoted on their posts. Personnel at all levels should be given an opportunity to improve their competence. Personnel from universities and research institutions should also be inducted at higher levels for a period of two to three years to give impetus to research and objectivity. This would, of course, need creation of suitable posts for them.

[*] Formerly Regional Advisor, ESCAP

## 4. Short-Term Economic Policy Requirements

Short-term policies require the collection and availability of data on weekly, monthly and quarterly bases. These relate to employment, production, sales, turnover, stock-holding, investment (expected and realised), exports and imports, prices, money supply, government revenue and expenditure.

The data collected on an annual basis should relate to employment, production, prices, gross output, input, investment, depreciation, separately for organised and unorganised sectors, etc.

A vast country like India requires statistics for decision-making on national, state and local levels. The collection, tabulation and analysis of data to be undertaken by various agencies should be laid down with an eye to their use in time. The data should relate to areas of concern at all levels such as functional groups, government, commerce and industry, and various social groups (including weaker and disadvantaged sections) classified by sex and age. They should separately provide timely information on all aspects of economic and social problems of children, especially girl children, and of women.

## 5. Data Gaps for Short-Term Policies

For business analysis, collection of data needs to be organised on monthly and quarterly turnover. For organised industry, surveys need to be initiated on planned and realised monthly, quarterly and yearly production. Data should also be collected on planned investment for the ensuing year, and expected and realised investment for each of the coming quarters. Similar investment data should also be organised for all infrastructure sectors.

## 6. Data on New Activities

With the liberalisation and globalisation of our economy, a large number of new activities/enterprises creating new types of income are emerging. Surveys on data collection in respect of these should be organised after evolving suitable methodology so as to produce usable data on earnings, incomes, etc.

## 7. Economic Data

The NSSO, a unique organisation established in the country, has been collecting data on various aspects of the economic and social life of the people. Although in most cases it has achieved great success, in economic fields it has not succeeded in all cases specially when it comes to data on income, savings and consumption expenditure of better-off classes in the society. Similarly, it has also failed to collect usable data on production, cost of production and income of enterprises engaged in transport, trade and economic services. In order to collect usable data on these items special methods and techniques are necessary. Methodological work and type studies can be best handled by research institutions and individual research workers.

## 8. National Accounts

National accounts have made rapid strides towards the compilation of various aggregates in great detail. Still, much needs to be done in the coming years. Statistics need to be organised on the basis of enterprise by industry. This implies that various categories of industries under the control of different enterprises should be presented in such a manner as to give various aggregates for all of them by nature of enterprises also. In the present-day world, with increasing competition, mergers and closures (bankruptcy) are regular phenomena. Hence data need to be collected continuously so as to be informed of the current status of enterprises.

Data on depreciation are obtained on the basis of estimated/assumed life of assets. These have not been found very realistic and hence there is a need to undertake censuses and sample surveys of various industries on the basis of their inventories of fixed assets. These require putting the inventory data in each enterprise/industry on computer and making them available to the statistical organisation on an annual basis. Wherever this is not being done, people in authority should be persuaded to do so. This work may call for a special project to be completed in five years with the help of a consultant or a research institution.

In the case of advance estimates (by which I mean estimates for the next financial year which are released in the last month of the current financial year), use is made of statistical information and statistical tools and models. The national income estimates of the ensuing year are primarily dependent not only on the internal situation but also on external factors. In the interdependent world, growth in the country is also conditioned by growth in other developed and developing countries. Such information is available in the publications of the OECD, IMF, World Bank, ESCAP, United Nations and a number of independent research organisations. This needs to be utilised to improve upon the estimates prepared on the basis of current methodology.

There has been a great improvement in the widening of the contents of the 1989-90 input-output table, making it all the more valuable for a large circle of users. This also provides a firm base for further work for the next input-output table in which, if possible, industrial categories should be expanded especially in the construction and tertiary sectors.

## 9. Enterprise Register

Currently, no enterprise register is maintained in the country. What is available at present is a classification into directory establishments and non-directory establishments and own-account workers by location, compiled once every ten years with the help of population census records, which are used as a frame for subsequent economic censuses and follow-up surveys.

Attention needs to be paid to the drawing up of an enterprise register for corporate enterprises, which should also include establishments under different enterprises by location. These should be continuously updated by keeping track of births and deaths of various enterprises. There are mergers, liquidation, bifurcation, etc. so that an annual survey should be based on the current knowledge of enterprise/establishment by employment, capital, etc. This needs creation of a unit in the CSO/NSSO which should constantly keep track of information by correspondence, newspaper reports and other means.

## 10. Statistical Authority

In India, the statistical authority is a decentralised one. This is what it should be. But problems arise when different ministries in the Central government and State statistical authorities do not follow the rules of the game, i.e. do not develop unbiased and reliable statistics according to a time-bound programme to meet various requirements. This seems to be happening in a large number of cases and hence the users in Central government, State governments, research institutions, individual research workers and people in general begin losing faith in a number of statistics. Use of unreliable and biased statistics have often made the administrators take wrong decisions which have cost the country dearly. It is, therefore, necessary to provide teeth to the statistical authority with a view to meeting genuine needs of the users and restoring credibility to the statistical system.

## 11. Research Institutions

Statistics is taught in most universities in India at the undergraduate as well as the post-graduate level. There is also the Indian Statistical Institute at Calcutta with its branches in Delhi and Bangalore. Quantitative research work has been done by all these institutions, but of late thrust on

quantitative work has been lacking. It is, therefore, important that an institute specially devoted to macro-economic research be established at an early date which should devote itself to quantitative research in economic and applied statistics at macro and micro levels. With a view to carrying out intensive research on regional aspects, regional institutes should also be established in various parts of the country during the coming decade.

## 12. Conclusion

The need of the hour is that statistics should be developed in all areas of concern and their reliability improved so that they can come close to the perception of the people. Various types of statistical tool for theoretical and applied statistics should be developed, research in various fields encouraged and strengthened through establishment of special posts in the CSO, financial assistance provided to established institutions in various universities, and new institutions established on macro-economic research. The data collected at the State and Central levels should be subjected to thorough scrutiny, and it should be seen that in the course of the next five years or so, national statistics emerges as aggregation of local-level and State-level statistics.

# MEASUREMENT OF POVERTY IN INDIA - A PLEA FOR STANDARDIZATION

## Shoutir Kishore Chatterjee[*]

For a developing country like India, where vast sections of the population suffer from chronic deprivation, the importance of measuring the extent of poverty at regular intervals cannot be over-emphasized. If, as a first step, we accept the poverty ratio (i.e. the proportion of the poor) as an index of poverty, then this means a time series of values of this ratio, suitably disaggregated, should be prepared and officially published by the Government just like it is done in the case of National Accounts, price and production index numbers, etc. In the regular official publications, however, one does not generally come across authentic statements of values of the poverty ratio. A strange kind of reticence on the part of the Government seems to prevail here!

True, the Planning Commission occasionally gives out values of the poverty ratio in its reports and documents (some of these find place occasionally in CSO booklets like *Selected Socio-economic Indicators*). These are the outcomes of studies carried out from time to time by teams of researchers and expert groups under its aegis. But if one sets together the figures put forth at different times by different institutions or researchers, one can discern wide fluctuations and blatant inconsistencies.

To illustrate, for the year 1977-78, the rural, urban and combined per cent poverty ratios for all-India were earlier (see *Selected Socio-economic Indicators for India*, CSO, 1989. Source: Planning Commission) reported as 51.2, 38.2 and 48.3, respectively, whereas in 1993, the Expert Group of the Planning Commission (Chairman: D.T. Lakdawala) in its Report gave for 1977-78 the revised figures 53.1, 47.4 and 51.8 in that order. Similarly at the time of the Seventh Five Year PLan, the Planning Commission gave 33.4, 20.1 and 29.9 as the all-India rural, urban and combined poverty ratio figures (in percentage form) for 1987-88, whereas the 1993 Expert Group reported the widely different values 39.1, 40.1 and 39.3 for the same ratios for the same year. More recently, early in 1996, the Planning Commission was reported to have given the combined poverty ratio for 1993-94 as 19 per cent and later in the same year, after the re-constitution of the Commission, a much higher percentage (36) was declared (*The Statesman*, December 26, 1996). Such inconsistencies and outrageous revisions have time and again drawn derisive comments in newspaper editorials; the commentators have often seen various motives behind the figures reported. Apart from the above, some serious researchers also have from time to time approached the problem of empirical measurement of poverty from their respective viewpoints. Their studies have produced new figures which have tended to confound the confusion. (See S.D. Tendulkar and L.R. Jain - *Economic and Political Weekly*, June 10, 1995 for one such recent piece of work.)

What is the reason behind such confusion? It is simply that there is no officially standardized and generally accepted method for computing the poverty ratio and different methods have been followed in different studies. After all, the official series of wholesale price index numbers and the different consumer price index numbers are generally accepted only because their methodologies (procedures for collection of price quotations, formulae for aggregation, weighting diagrams, etc.) have been standardized. It is high time that a consensus is reached by all concerned

[*]CSIR Emeritus Scientist attached to the Department of Statistics, University of Calcutta

regarding a standard method of computation of poverty ratios. Once such a method is evolved and agreed upon at the official level, it should be adhered to for medium-range periods.

The first step in evolving a method for poverty measurement is to define 'a poor person'. In 1979 the Task Force (Chairman: Y.K.Alagh) appointed by the Planning Commission gave a definition in terms of minimum calorie intake, which has been generally accepted in most poverty studies in India during recent years. According to it, a person in rural (urban) areas would be considered poor if his or her daily calorie intake is less than 2400 (2100) units. On the basis of well-defined baskets of food-items appropriate at the all-India level, the monthly per capita total expenditures (MPCTEs) whose components for food would ensure these minimum consumptions turn out to be Rs.49.09 (56.64) at the all-India prices of 1973-74 (the anchor year). These minimum MPCTEs or poverty norms represent the official 'poverty lines'. At any time for any section (rural or urban) of the population, if one knows the frequency distributions of MPCTE, one can determine the poverty ratio as the relative frequency to the left of a cut-off point which is equal to an appropriately deflated value of the corresponding poverty norm. For the country as a whole, the NSSO has been conducting full-scale household consumer expenditure surveys roughly at quinquennial intervals (1977-78, 1983, 1987-88, 1993-94). (Besides, since 1986-87 (42nd round), similar surveys on a smaller scale are being conducted in the intermediate years as well.) From these surveys one gets the frequency distributions of MPCTE for the rural and urban areas of every State (by this, we mean Union Territories also) as well as the entire country. Most studies utilize these NSS-generated MPCTE-distributions for computing poverty ratios. But there are different modes of utilization of these, representing different methods of computation.

Broadly, there are two approaches here:

1. based on the MPCTE-distributions derived from NSS household consumer expenditure data alone,
2. based on NSS MPCTE-distributions revised in the light of the Private Final Consumption Expenditure (PFCE) estimates of National accounts prepared by the CSO.

Of these, the first approach has been favoured generally in most studies on the measurement of poverty in India; the Planning Commission Expert Group, 1993, recommended that only this approach should be followed. The second approach was adopted by the Planning Commission in the 80's in the Seventh Plan document and it led to the reporting of inordinately low values of the poverty ratio (did the Commission revert to the same approach early in 1996?). Serious misgivings have been expressed about the validity of this approach. However, before we dispose of this approach, let us first make a few observations on its shortcomings.

To describe the essence of the second approach, let $R_0(x)$, $U_0(x)$ and $F_0(x)$ denote the cumulative (relative frequency) distribution functions (cdf's) of MPCTE for rural areas, urban areas and the whole of the country based on the NSS household consumer expenditure survey in a reference year (year 0). If $p_0$ denotes the proportion of the population living in rural areas in that year, clearly

$$F_0(x) = p_0 R_0(x) + (1 - p_0) U_0(x).$$

Let $c_0$ denote the mean MPCTE for the country for the year based on $F_0$. For a nearby subsequent year t, let $c_t'$ denote the per capita figure obtained by dividing the total PFCE of National Accounts by the corresponding population size. Writing $\rho_t = c_0 / c_t'$, it is postulated that the true distributions of MPCTE in the year t in rural and urban areas would be given by the cdf's $R_0(\rho_t x)$ and $U_0(\rho_t x)$,

respectively. (In the Indian setting, we always have $\rho_t < 1$, so that such an adjustment results in moving the distributions stochastically to the right.) The argument for this adjustment presumably is that if we write

$$F_t(x) = p_0 \, R_0(\rho_t \, , x) + (1 - p_0) \, U_0(\rho_t \, , x),$$

then the mean MPCTE for the country based on $F_t(x)$ would be $c_t'$. If $r_t$ and $u_t$ are the updated rural and urban poverty norms for the year t (obtained by suitably deflating the poverty norms for an anchor year, the rural and urban poverty ratios would be obtained as $R_0(\rho_t \, , r_t)$ and $U_0(\rho_t \, , u_t)$. It has been pointed out (see, e.g., B.S. Minhas (1988), *Sankhyā*, Ser.B, Vol. 50, Part 3 supplement) that because of basic differences in concepts and the time lag between production and consumption, $c_t'$ and the counterpart of $c_0$ for the year t are not comparable. But even leaving aside the question of empirical comparability, from the above it is clear that $R_0(\rho_t \, , x)$ and $U_0(\rho_t \, , x)$ would represent the correct rural and urban MPCTE distributions for time t provided between the years 0 and t, the MPCTE of every person changes by a constant factor, which is the same for the rural and the urban sectors and the rural-urban composition of the population remains the same. Such sweeping assumptions, needless to say, are totally unrealistic. The situation becomes even worse if we try to find State-wise disaggregated values of rural and urban poverty ratios from the corresponding MPCTE-distributions for year 0 of each State by using the same (all-India) adjustment factor $\rho_t$ in all cases.

To describe the essence of the approach based on MPCTE-distributions alone, we confine ourselves to the rural sector (the treatment of the urban sector will be along parallel lines). Let $r_a$ denote the all-India rural poverty norm with reference to an anchor year (year a). (As mentioned earlier, most recent studies take 1973-74 as the anchor year and $r_a$ = Rs 49.09). In any year t in which a full-scale NSS household consumer expenditure survey is held, let $R_{ts}(x)$ and $R_t(x)$ denote the cdf of the distribution for the s-th State (s = 1,2...,S, say) and the entire country. If the proportion of the rural population of the country living in the s-th State is $W_{ts}$, we have

$$R_t(x) = \Sigma_s \, W_{ts} \, R_{ts} \, (x).$$

One can get the all-India poverty ratio by using the counterpart of $r_a$ for the t-th year in $R_t(x)$. But because there are significant differences in the price-levels among the States and these levels move with time differently in different States and also because Statewise figures are of interest, a disaggregated approach is preferred here. This means that in the year t, for the s-th State we have to find the counterpart $r_{ts}$ of $r_a$ and compute $R_{ts}(r_{ts})$ as the poverty ratio, s = 1,2...,S. The all-India figure may then be obtained as $\Sigma_s \, W_{ts} \, R_t(r_{ts})$.

If we have spatial price indices $J_s$, representing the price level in the s-th State relative to the all-India level (taken as 1), for the anchor year a, the State's poverty-norm in its own anchor-year prices is $r_{as} = r_a \, J_s$. If further, the temporal price index of year t relative to the price level (taken as 1) of year a is $I_{ts}$ for State s, we can compute $r_{ts} = r_{as}I_{ts} = r_a \, J_s I_{ts}$ as the poverty norm of State s in the year t.

Thus, we can get the poverty ratios for any year t for which the cdf's $R_{ts}(x)$, s = 1,2..., S, are available. Fairly reliable MPCTE-distributions are available State-wise in those years (roughly at quinquennial intervals) in which full-scale consumer expenditure surveys are conducted. Once the spatial and temporal price indices $J_s$ and $I_{ts}$ are chosen, computation of poverty indices is straightforward for these survey years. Diversity in the results obtained by different workers arises

because of differences in the indices $J_s$ and $I_{ts}$ used. As regards $J_s$, indices based on weighting diagrams representing the consumption patterns of roughly the second and third quintile groups from the bottom (which generally straddle the poverty lines) of the rural and urban populations have been recommended [see B.S.Minhas and L.R.Jain (1990), *Tech Report 8915*, ISI, New Delhi, which contains other references]. Regarding $I_{ts}$ various choices, such as the National Income consumption expenditure deflator, WPI, for the rural sector CPI for agricultural labourers and Special CPI based on consumption habits of middle groups, and for the urban sector CPI for industrial workers and non-manual employees and special CPI for middle groups [see Minhas and Jain (1990)], have been made and recommended. As regards the years intermediate between successive survey years, MPCTE-distributions based on smaller (thin) samples are now available. Doubts have been expressed [see Tendulkar and Jain (1995)] as to whether the State-wise MPCTE-distributions for these years are reliable enough to permit a disaggregated approach as above. Direct computation of all-India poverty ratios and some indirect method to get the State figures may be appropriate for the intermediate years.

As the above account shows, to standardize the procedure for the computation of State-wise rural and urban poverty ratios on a regular basis, we have to make definite choices as regards the arbitrary elements in the different steps of the procedure.

Firstly, we have to fix the poverty norms for each State with reference to a recent year. This can be done by first choosing a suitable basket of goods at the all-India level to ensure the normative minimum calorie intake. Discussing again with reference to the rural sectors, we get the corresponding MPCTE $r_a$ as the all-India poverty norm. On the basis of suitable weighting diagrams, the State-wise spatial indices $J_s$ for the anchor year can be determined and hence the poverty norm $r_{as} = r_a J_s$ for the s-th State can be found. However, since the consumption pattern varies over the States, it may be more realistic to take different baskets of goods ensuring the same minimum calorie intake in different States and derive the $r_{as}$ values directly. A similar procedure may be followed for the urban sector. As regards the calorie norms, taking some additional levels apart from the Task Force levels of 2400 and 2100 units would mitigate to some extent the arbitrariness of the poverty ratio as an index of poverty.

After the State-wise poverty norms have been found for the anchor year, the same for any other year t can be determined by using suitable consumer price indices of the type $I_{ts}$ mentioned earlier. Series of such indices for the rural and urban areas of each State may be prepared on the basis of weighting diagrams reflecting the consumption patterns of those located around the borderline of poverty. Once fixed, the weighting diagrams should be adhered to for reasonable periods of time. The new series of CPIs should be regularly included in official publications.

Lastly, for those intermediate years in which full-scale household consumer expenditure surveys are not conducted by the NSSO, some reasonable method of estimation of State-wise poverty ratios should be decided upon. Currently, in the NSS, in any such intermediate year a small number of households is selected from every selected first-stage unit (fsu) (village/urban block) for canvassing the consumer expenditure schedule. If the survey procedure could be modified so that the thin sample of households taken in an intermediate year is a sub-sample of the regular sample of households selected in the earlier full-scale survey year (this would necessitate that the fsu's of the thin sample form a sub-sample of the fsu's selected in the earlier year), for every household in the thin sample we would get the consumption expenditures for the current year as well as the earlier year. The idea of two-phase sampling as extended to left tail-area estimation can then be employed to get reliable estimates of State-wise poverty ratios in the intermediate years by utilizing the

MPCTE-distribution based on the larger sample of the earlier year. This would, of course, entail some complication in the survey procedure.

The bottom line is that the National Advisory Board on Statistics and the Government should take a plunge in standardizing the procedure of poverty measurement and publishing regular series of all-India and State-specific poverty indices. True, poverty measurement is a sensitive issue, but so are the issues of estimation of GDP and its components and the fixation of CPIs. Objectivity and transparency in the procedure and boldness in execution will exonerate statisticians and the authorities from charges of 'mendacity', 'doctoring the data' and 'playing with poverty' and will have a long-term salutary effect on keeping the efforts of the Government in the right gear, whoever may be at the helm of affairs.

---

*Editors' note*: Using NSSO data, Dubey and Gangopadhyay ("Counting the poor", *Sarvekshana*, February 1998) have recently done some useful studies on poverty relevant to the present context.

# VIEWING NSSO FROM OUTSIDE AND INSIDE

## Arijit Chaudhuri[*]

In the first part of this article I intend to present the views on some NSSO activities that I held when I was totally an outsider. In the second part, I shall express some of the modifications of these views developed since I ceased to be so in recent times.

The NSSO appears to be a highly conservative organisation very zealously sticking to certain age-old programmes and practices. The subject coverages in annual or quinquennial cycles follow essentially unaltered patterns. The questionnaires are highly structured and detailed. In fact, they are so elaborate that they are bulky to the extent of appearing repulsive to their addressees. The sampling designs proverbially involve stratification and selection of units at two stages. Sampling is done separately for rural and urban India. The urban primaries are chosen circular systematically or by the simple random without replacement method and the urban secondaries also by either of the same procedures. The rural primaries, on the other hand, are selected PPS circular systematically or by PPS with replacement while the rural secondaries are chosen in the same way as the urban secondaries. Another ancient habit is the insufferably inordinate procrastination in publication of the findings, though of late there are clear signs of improvement in this regard. Nevertheless, when they emerge, the coverage becomes huge in terms of the numbers of tables with numerous titles. But one palpable consequence of a lengthy questionnaire is the multiplicity of unoccupied cells in most of the two-way tables. Another characteristic of NSSO reports is the absence of estimated standard errors with a claimer that they are calculated using the celebrated 'half-sampling technique' but not recorded to save publication space. There has never been any lack of appreciation of the NSSO's organizational skill and acumen in implementing the unenviably daunting task of generating the grassroot level raw data on the nation's gigantic scale at a highly desirable level of quality.

A major criticism against the NSSO has been its policy of denying the outsiders access to the basic questionnaire-wise data it procures. A second criticism validly levelled against it is the requirement that any publishable paper using NSSO data must first be subject to the review process of its own official journal called *Sarvekshana*. However, these policies have recently been revised. Being a student of survey sampling, I have been nurturing a long-standing grievance that no innovative techniques of survey sampling are being tried by the NSSO over the years, though it started as a pioneer in the world in implementing large-scale surveys with many novel ideas initiated by Professor Mahalanobis, the father of modern statistics in India. Let me be slightly specific about such flat charges.

It is generally believed that the effects of non-sampling errors on large-scale survey results far outweigh those of sampling errors. So, a well-motivated user of NSSO reports may legitimately ask what scientific steps are being taken to eradicate any possible adverse influences of non-sampling errors on the results the NSSO releases. Satisfactory answers are hard to come by. Outside India, techniques are emerging fast to tackle both 'unit' and 'item' non-responses by dint of 'imputation' and 'weighting adjustments' and other requisite remedial measures. Diverse methods like 'balanced half-sampling', 'balanced repeated replication', 'jack-knifing' and 'bootstraps' are commonly employed devices in estimating standard errors and setting confidence limits for

---

[*] Professor, Indian Statistical Institute, Calcutta and Member of the Governing Council, NSSO.

estimand parameters. The NSSO so far has not cared for these. Analytic studies are avoided as a rule. Only descriptive statistics are developed here. For example, correlation and regression analyses are not reported by the NSSO. To embark upon such programmes by the users of NSSO reports again is not easy in the absence of more basic data in them.

To start the second part of the narration, I must admit my awareness that the NSSO lacks resources to have a supporting research wing with the capability to continually keep abreast of the rapidly changing world and a well-equipped capacity for critically assessing the possibilities of any shortcomings in its own omissions and commissions and endeavouring to forge ahead with suggestions and recommendations for innovative and corrective measures. It is not in a position to launch a series of seminars to examine its own rôles with exposures to the scrutiny by knowledgeable outside participants. It is not well-equipped to have its own journal *Sarvekshana* acquire the status of a readable scientific journal with original contributions on methodological issues. It is deficient in men, machines and money. Some hot topics developed in the post-Mahalanobis era, namely, 'small area estimation', 'randomized response', 'network sampling', 'adaptive sampling', 'remote sensing', etc., apparently are not in the common vocabulary of those who guide the courses of activities of the NSSO at the present moment. It is commonplace knowledge that small area estimation is a crying need in India in connection with decentralised planning, and the advantages of applying techniques already developed could be immense. I am greatly tempted to take the advantage of this piece of submission to offer a proposal to initiate a few projects to do some exercises covering the above techniques to take care of some national or regional concrete problem areas. If called upon, I offer myself as a participant in any of these to the extent I may be permitted by the organization to which I am affiliated.

# A BRIEF NOTE ON SAMPLE SURVEYS

## V. P. Godambe[*]

Development of Sample Surveys in India over the last fifty years has been spectacular. New techniques of sampling and methods of estimation have led to increased efficiencies.

Sample surveys generally aim at providing estimates of population parameters like totals (means) of various items. Computation of errors of the estimates is considerably simplified for interpenetrating sampling designs. Though these 'errors' provide a measure of accuracy of the estimation, confidence intervals based on them are often inaccurate, particularly for small sample sizes. This is because the distribution of the quantity {(estimate-parameter)/error} cannot be approximated by the standard normal distribution. On the other hand, if in the just mentioned 'quantity' the 'error' in the denominator is replaced by a sample-based parametric function which approximates the square root of the variance of (estimate-parameter) in the numerator, the 'resulting' quantity has nearly standard normal distribution. Hence this alternative quantity provides more accurate confidence intervals. Computation of these confidence intervals is naturally more involved than the former ones. But with the availability of fast computers, this need not be a big problem. For details we refer to the paper "A new look at confidence intervals in survey sampling" (1998) by the author.

At the moment, in survey practice, confidence intervals are seldom computed; instead, coefficient of variation is computed. But, in the near future, when use of confidence intervals in survey practice would be common, the methodology suggested above would come in handy.

---

**Reference**

Godambe V. P. (1998). "A new look at confidence intervals and survey sampling". *Working paper 98-02*, Dept. Stat. Actuarial Sc., University of Waterloo.

---

[*] Professor of Statistics, University of Waterloo, Canada.

# A NOTE ON THE MEASUREMENT OF INDUSTRIAL PRODUCTION

## Kishori Lal[*]

Measurement of Industrial Production has been there for a long time, much longer than the more widely known and reported of Gross Domestic Product (GDP) of the System of National Accounts. Industrial Production is regularly published by the Organization for Economic Co-operation and Development (OECD) as one of the main economic indicators. The world-famous weekly British magazine, *The Economist*, publishes, in every issue, industrial production of not only the advanced industrial countries but of all the 25 emerging economies of the world. The emerging countries list includes India, China, Russia, Taiwan, Turkey and others. There is no doubt about the importance attached to this concept and its measurement. Given its age and its importance, one would expect that this concept is well defined and that it follows internationally accepted guidelines. This note would try to test this hypothesis.

## 1. Boundary of Industrial Production

Industrial Production includes the following three divisions of standard Industrial Classification in Canada : Mining, Manufacturing and Utilities. Most (but not all) member countries of the OECD include the same set of industrial divisions. Utilities such as Electric Power and Gas Distribution are typically the ones ignored in some countries. For any inter-country or even inter-provincial comparison, the first check needs to be made regarding the units included underlying the concept at hand.

After one has assured oneself about the coverage of units, the next big question relates to the meaning of production. Does production mean gross output like cars in a car factory or value added in the production of cars? There is no international agreement on this point, not even among the OECD member countries. Note that while the value added is an unduplicated concept, the gross output includes duplication. Further, the amount of duplication depends upon how the industrial structure in a given jurisdiction is integrated. The more it is integrated, the less the gross output, the limit being full integration or full consolidation which would yield gross output equal to value added only. This is a bit complicated and would require further elaboration later.

Industrial Production is a volume concept, not a current price concept. There is even less agreement at the international level regarding the precise formulation of constant price Industrial Production. Industrial Production at constant prices and the Index of Industrial Production will be different (and the difference would not always be small) depending upon the base one chooses, how often one changes the base year and the base weights one chooses. For example, one might calculate the index of industrial production in 1998 over 1997 using 1997 weights or 1992 weights or even 1981 weights. Each of these formulations would provide different growth rates for 1998 over 1997. Are the weights derived from gross output or value added? This again would result in differences in the 1998 growth rate. Is the statistical agency using a single deflator regime or the preferred double deflation technique?

## 2. Units Included in Industrial Production

As noted above, the three divisions of Standard Industrial Classification - Mining, Manufacturing and Utilities - are the norm for the industrial production statistics. The underlying detail at three or four digit industries forming these divisions will always be different amongst the

[*] Director General, System of National Accounts Statistics, Canada

various countries. It is the industrial detail which is really the more interesting aspect of this phenomenon. In Canada, we estimate for about 160 separate industries for industrial production and aggregate them into 100 industries for public release, all at monthly frequency. For inter-temporal comparisons, one must have a consistent definition for each and all aggregations of industries for the entire period under consideration. We have just released the above detail for the period 1961 to date and, of course, at the monthly, quarterly and annual detail. It was quite a time-consuming exercise to diligently classify the detailed industries from the various vintages of Standard Industrial Classifications used over the last four decades in Canada.

### 3. Definition of Production

Production for any aggregation is not a well-defined term as it could include any level of duplication. Let us assume that there are two enterprises, one in Delhi and the other in Calcutta, and both produce cars. The two enterprises are identical in technology and scale. But the enterprise in Calcutta is organized in two establishments - the first producing parts and the other producing cars using parts from unit 1. The enterprise in Delhi has both these units combined into one; other than that, there is no difference. The establishment producing parts in Calcutta buys Rs.300 million of raw materials to produce parts which it sells to the car producing establishment for Rs.700 million, thus adding a value of Rs.400 million. The car producing establishment in Calcutta uses those parts and produces cars worth Rs. 1,500 million, thus adding the value of Rs.800 million. The value added at the two establishments in Calcutta equals Rs.1,200 million (400 in the parts factory and 800 in the car factory). The Delhi enterprise is an integrated unit; it buys raw materials for Rs.300 million, makes parts and uses them for cars and produces cars worth Rs.1,500 million, an amount identical to the Calcutta establishment. The value added in Delhi is Rs.1,200 (1,500 less 300) million. Both places produce an identical value added but the gross production is different. Value of production in Calcutta is 2,200 (Rs. million) (1,500 for cars and 700 for parts). Value of production in Delhi is 1,500 (Rs. million), all cars, because the parts establishment is not separated. Value of weight of industrial production is 3,700 in the above example if one picks up the reports from the enterprises; but only 3,000 if one ignores the administrative difference at the two locations and only concentrates on the finished product; and only 2,400 if one counts the value added from these two locations. The internal shuffling does not change the value added but does change the so-called value of production.

### 4. Value Added Weights

A very important industrial country in Europe used to publish a different value of production for every different aggregation of statistics. It was quite confusing. Let me rephrase that the only concept which would produce the same value of output, irrespective of the level of aggregation, is value added. There is a big advantage to use value added as a weight in aggregating the detail for industrial production. At Statistics Canada, we use the value added weights for aggregating each of the 150 separate industries which form part of Canadian industrial production. Another advantage of using value added weights for industrial production is that its importance can then easily be compared with other sectors of the economy such as the services sector or government sector or agriculture. An even more important benefit is that the industrial production program becomes a part of the GDP by industry program. It is then automatically produced in an integrated set of statistics, giving this program a higher quality rating and placing it in a better analytical framework.

### 5. Deflation of Industrial Production

At the outset, it is important to note that measurement of industrial production at constant prices or, for that matter, real output of any sector or for the total economy is essentially the construction of a model or logical abstraction of actual transactions. "Volume" or "quantum" is always estimated with reference to time and weights. As we are all familiar with the index number

problems associated with changing weights and aggregation, the estimated quantum measure at any level of aggregation is not an unambiguous phenomenon even if one has the most reliable value, price and quantity statistics in detail. This nature of deflation applies to goods, services and industries, whatever methods are used. Value added or industrial production at constant price can be derived by the preferred double deflation technique. Such constant price value added may be extrapolated using constant price gross output, physical quantities of gross output, constant price intermediate inputs, employment, etc. A connected issue is how often one should rebase the constant price series and then how to link the earlier period growth rates when the base year changes.

## 6. Double Deflation

In principle, any commodity forming part of industrial production can be deflated using price indices or quantities produced as there are units of measurement available. But there is no unit of measurement of value added for industrial production or value added for other sectors of the economy or even the entire economy. Hence it cannot be directly deflated. One needs some other technique to convert current price GDP into constant price GDP. The internationally recommended preferred technique is double deflation. In this approach, one deflates commodity outputs of an industry at basic prices, its intermediate uses of commodity inputs at basic prices as well as taxes on products and other taxes on production, net of subsidies. The difference between the deflated values of outputs and the total of commodity inputs and net taxes on products and production equals GDP at factor cost in constant prices. This approach, called double deflation approach, satisfies the requirement of an identity between GDP income and expenditure based estimates in constant prices. However, there are certain important hazards in using double deflation for deriving the GDP of an industry whose value added represents a small proportion of total gross output. We will come back to this issue later.

Note that the double deflation approach assumes the availability of Input and Output tables or Make and Use tables or, in the 1993 SNA terminology, the Supply and Use tables. We strongly support the 1993 SNA recommendation to use the double deflation approach to calculate constant price GDP or GDP of any sector such as industrial production. The 1993 SNA states: "... the supply and use tables are the most complete consistent framework for constant price estimation and provide: a) interdependent measures of prices and volumes; b) an important check on the numerical consistency and reliability of the entire set of such measures, interlinking values at constant and current prices, value and volume indexes and deflators" (paragraph 15.161). "Constant price measures for gross value added are possible in the input-output framework by using the double deflation method, as the difference between: a) the value of output deflated by a price index of outputs; 2) the value of intermediate consumption deflated by a price index for these inputs" (paragraph 15.162).

The double deflation method in the context of inputs and outputs at basic prices is a very efficient technique. A similar recommendation was made earlier in the *UN Manual on National Accounts at Constant Prices*, Series M, No. 64, New York, 1979 (in short, *UN Constant Price Manual*). Let me further discuss this issue, using the Canadian practice. The Canadian practice conforms to the recommendation in the *UN Constant Price Manual*. The *UN Constant Price Manual* states: "In an ideal world real product by kind of activity would always be derived from an input-output table by double deflation" (p.55). In Canada, the input-output tables form the core of the production accounts. Input-output tables in full detail are produced annually both in current and constant prices. Real output by industry and constant price Industrial Production are produced using the preferred double deflation approach.

I noted above that there are certain hazards using double deflation for those industries whose value added component makes a very small share of total output. For such an industry, GDP estimated by double deflation might be erratic, because small shifts in the relative prices of intermediate inputs and gross output could translate into big shifts in the resultant value added at

constant prices. Here, the *UN Constant Price Manual* guidelines are not entirely satisfactory. They state: "The solution to this problem, however, may simply be to consolidate industries with very small ratios of value added to gross output with related industries at earlier or later stages of production. In other words, the problem of instability may be solved by aggregation into larger units whose values added are large enough in relation to gross output not to be too sensitive to the effects of changes in prices or technology" (p.53). In one year, value added in one industry may be erratic, but in the next, a different industry might suffer. *Ad hoc* aggregation into large units would disturb the continuity of time series. Thus, one needs additional guidelines. In the Canadian System of National Accounts (CSNA), we have solved this problem as follows: Values added are combined as suggested by the *UN Constant Price Manual*. The combined value added is redistributed using gross output or any other indicator as a proxy, but the combined value for a given sub-aggregate remains unchanged. Without this restriction, the above-noted GDP identity requirement will not be satisfied. These comments on the *UN Constant Price Manual* apply equally to the 1993 SNA recommendation on double deflation.

## 7. Chain Indices

The 1993 SNA provides a framework within which an integrated set of price and volume measures can be compiled which are conceptually consistent and analytically useful. The 1968 UN SNA provided a summary discussion on price and volume comparisons, but it was limited to the production accounts. More detailed guidelines were issued in 1979 by the United Nations in the *Manual on National Accounts at Constant Prices*, Series M, No. 64, United Nations, New York, 1979. The 1993 SNA differs from both the 1968 UN SNA and the 1979 *UN Constant Price Manual* in its recommendation to use chain indices.

It is self-evident that the more remote the base becomes, the less relevant are its prices for purposes of deflating the value of current flows of goods and services. When the base year is changed, there are two ways in which the series on either side of the new base year may be linked. The first method is to revalue not only the series for all years subsequent to the new base year at new prices, but also the series for all the years preceding the new base year. The second method is to leave the data for all years up to and including the new base year unchanged and simply use the new base year prices for valuing all flows of goods and services from the new base year onwards. In order to develop an unbroken series expanding on either side of the new base, one would simply link the series through a chain index. We reject the first method in favour of the second on conceptual grounds. A new base year is required because the old one is not relevant any more: thus, one cannot effectively argue for the first method. Both the 1968 *UN SNA* and the 1979 *UN Constant Price Manual* opted for the first method whereas the 1993 SNA recommends the use of chain indices for linking the series.

The 1993 SNA recommends: "In general, the constant price series should not be allowed to run for more than five, at the most, ten years without rebasing" (paragraph 16.76). It further recommends: "The preferred measure of year to year movements of GDP volume is a Fisher volume index, changes over longer periods being obtained by chaining: that is, by cumulating the year to year movements" (paragraph 16.73a). This is a very significant recommendation. Its implications for all countries, particularly developing countries, are very significant. There is a lot of debate in the media, in the legislatures and in the scholarly journals about growth rates. A fair amount of this debate is meaningless, because the growth rates measured with reference to fixed weights established a decade or so earlier are misleading. As the relative prices have become very volatile, fixed weight growth rates are not reliable. A five per cent growth rate could easily be three or four per cent when the base year is changed. A few years ago, Allan H. Young, the Deputy Director of the Bureau of Economic Analysis, US Department of Commerce, published a paper, "Alternative measures of change in real output and prices", *Survey of Current Business*, April 1992, in which he compared (constant price) average annual rate of change in the manufacturing sector (a most significant part of industrial production) of the United States, using three different fixed weight periods. The average annual rate of change for total manufacturing for the full decade 1977-87 was

4.7% using 1977 weights, 2.6% using 1982 weights and a mere 1.6% using 1987 weights. These are horrendous differences for any analysis. As both the overall rate of inflation and the relative price changes are much higher in India than they are in the United States, the growth rate changes due to a shift in base year might be even more dramatic. To get a sensible reading of a growth rate, one must change the fixed weight to as current a period as possible, or preferably use chain indices.

## 8. Additivity Problem

When the series preceding the new base year is chained to the series succeeding the new base year, "the problem that emerges ... is that the constant price values for the components do not add up to the constant price values of the aggregates after the series have linked" (*1993 SNA*, paragraph 16.37).

Addressing the additivity problem, the *1993 SNA* argues in favour of publishing linked, rebased data " ... without adjustment leaving it to users to decide whether, or how, to deal with the resulting discrepancies" (paragraph 16.59). This is the approach followed in the Canadian system but with a modification. In the CSNA's fixed weight indices for final expenditure (rebased and linked every five years) additional "adjusting entry" series are shown for each published aggregate, calculated as the difference between the linked aggregate and the sum of its published, linked components. Our approach has two advantages: it alerts users to the problem by displaying explicitly the extent of non-additivity, and it reassures them that no elements are missing from the accounting system.

## 9. Extrapolation of the Base Year Value Added

Even in many advanced countries, the input-output tables are not produced with an annual frequency. Even when the tables are produced annually as in Canada, the tables are produced with a time lag of about three years. Thus, there is a need to project industrial production to the current period using techniques other than double deflation. One can extrapolate value added using constant price gross output, physical quantities of gross output, constant price intermediate inputs, employment, etc. Should one prefer deflated output to physical quantities of gross output as a projector of base year value added? The *UN Manual* recommendation is: "The value at constant prices of the goods available in both years should be obtained by deflating the current year value by a price index, rather than by extrapolating the base year value by a volume index..... The justification for this recommendation is that price relatives generally display less variation than quantity relatives. The range of variation of quantity relatives may be anywhere from zero to infinity whereas that for price relatives is much narrower" (pp. 46-47).

If in both the base year and the current year (a) the values of all relevant transactions, (b) the quantities of all goods and services and (c) the prices of all goods and services are recorded, then both choices give the same results and one need not worry about which method one uses. In reality, such a complete recording does not exist, i.e., a statistician usually faces the problem of incomplete price and quantity information. Statisticians may still prefer quantities to prices if in their judgement the quantity information is less incomplete than the price information.

## 10. Concluding Remarks

The value of the measure of Industrial Production is greatly enhanced when it is situated in a time series context and is comparable across jurisdictions. This leads us to recommend the following: weights for industrial production should be value-added weights; the preferred double deflation approach should be used for constant price value added; and the base year for prices should be as current as possible, preferably no later than five years. Any national statistical agency must always take into account its own unique institutional features but must be aware of the cost of not producing the highest quality measure for such important statistics as statistics of industrial production.

# IMPROVING THE INDIAN STATISTICAL SYSTEM: SOME SUGGESTIONS

## G. C. Manna

### �301. Introduction

India, as on today, has developed a sound statistical system so far as collection, compilation and dissemination of data relating to socio-economic (including demographic), agricultural and industrial statistics are concerned. Various agencies like the Departments of the Central/State governments, institutions/autonomous bodies and non-government organisations (NGOs) are involved in this huge task of creation of database. Among the different agencies, the major sources which collect/compile huge volumes of data and also disseminate them more or less regularly at the national level, State/ Union Territory level or even at the sub-State level, are the Office of the Registrar General of India (RGI); the Ministry of Agriculture, Government of India; the Central Statistical Organisation (CSO) and the National Sample Survey Organisation (NSSO). The office of the RGI has been compiling statistics on various aspects of economic and social life of the Indian population based on the data collected through decennial censuses. Similarly, the DES and DAHD of the Ministry of Agriculture, *inter alia*, generate information on agricultural statistics and livestock population, respectively, based on the data collected through periodic Agricultural Censuses and Livestock Censuses. The CSO is responsible for coordination of statistical activities in the country as well as for evolving and maintaining statistical standards. Its activities include compilation and release of National Accounts Statistics, conducting the Annual Survey of Industries (ASI) as well as Economic Censuses and their follow-up surveys, compilation of the Index of Industrial Production and Consumer Price Indices for Urban Non-Manual Employees besides organising training in official statistics. The NSSO, through its nation-wide, large-scale, continuing sample surveys, conducted in the form of successive rounds, collects and disseminates information on diverse subject areas. It is also responsible for undertaking the field work for the ASI, updation of the urban frames through Urban Frame Survey (UFS) for the entire urban area of the country and performing sample checks on area enumeration and crop estimation surveys.

### 2. Reviewing the System

From what is described above, it is apparent how diversified and how voluminous are the data collected by the different agencies. A time has perhaps come to take stock of the quantum of data collected/being collected and to examine their usefulness. In case collection of data on a particular subject is required to be continued, it needs to be seen whether any change in the methodology is called for in order to remove the existing deficiencies. Based on my personal experience and association with the statistical activities, a few suggestions are given in the subsequent paragraphs. It is felt that the suggestions, if accepted, may help in improving the statistical system of the country to a great extent.

### 3. Suggestions

**Sampling Frame and its Updation** The list of villages provided by the latest decennial population census is used as the sampling frame for selection of sample villages in the rural areas by the survey

---

* Joint Director, SDRD, NSSO, Calcutta. Views expressed in this article are those of the author and not necessarily of the organisation to which he belongs.

agencies undertaking household or industrial enquiries. Generally, a PPS sample of villages is selected/recommended for the purpose of drawing of sample villages. This results in selection of a very large number of *larger* villages. Before canvassing the questionnaires, a frame of households is first prepared in the selected villages from which a sample of households (or *enterprises*) is drawn for detailed enquiries. As listing of all households/enterprises in the *larger* villages requires a huge amount of listing time, such villages are generally segmented, from which one or two segments are selected by random sampling and the listing is restricted within the selected segment(s). Finally, a sample of households (or enterprises) is selected from the frame of households/enterprises prepared within the selected segment(s).

The above procedure of sample selection introduces another stage of sampling error and thus further affects the quality of the survey data. To overcome this problem, it is strongly felt that the *larger* villages (a suitable cut-off point in terms of population of the village, say 800, can be decided so that the villages having populations in excess of the prescribed cut-off point can be termed as the *larger* villages) are segmented into a number of *permanent* segments during the population census. The existing method of dividing the larger villages into a number of enumeration blocks (EBs) with more or less equal population content and repeating the procedure (i.e. changing the EB boundaries) in different censuses may be modified. It is not necessary that the segments (presently called the EBs) should have equal population. What is most important is that the segments should be compact areas with permanent, well-defined, identifiable boundaries (viz. existence of roads, railway lines, etc.) and they are not changed in the subsequent censuses once they are formed (say, during the census of 2001). The population of each such segment might not exceed certain limit so that the segments (to be henceforth called EBs) can be directly selected as the first-stage units (by PPS sampling scheme or otherwise) for *larger* villages.

The village/EB maps should be prepared with adequate care so that the "permanent boundary" particulars are clearly indicated in the maps. In this stage of work, assistance/guidance of NSSO officials engaged in updation of Urban Frame Survey (UFS) blocks may be taken, if required, to ensure perfectness of the work.

An exactly similar procedure is suggested for formation of EBs and preparation of EB maps in the towns and cities also.

Certain additional auxiliary information at the village and EB (both rural and urban) levels, like availability of certain facilities (e.g. hospitals, nursing homes, homes for the aged, rehabilitation centres for the disabled, etc.), type of area (i.e. residential area, bazar area, industrial area, slum area, etc., as collected in the UFS) along with the information on population and number of households, may also be collected during the population census. The same can be utilised for stratification and selection of villages/EBs in the sample surveys, if required, depending upon the subjects of enquiry.

The list of villages and EBs (computerised) along with their maps may be made available to the survey organisations including their field offices (like the offices of FOD and SSBs) for using the same at the time of sample surveys.

The Economic Censuses (ECs) conducted along with the house-listing operations of the population censuses may continue to give, *inter alia*, the count of number of enterprises (mainly the non-agricultural enterprises) of different types at the EB level so that the information can be utilised for drawing of samples (EBs) for the follow-up surveys of economic censuses. In view of the rapid growth/decline of enterprises, one more EC may be conducted exactly at the mid-point of the period

between two population censuses. During this EC, the frame of EBs might be updated by considering the phenomenon of de-classification of areas (i.e. rural urban or *vice versa* in a few cases). However, the boundaries of the already formed EBs (formed during the bench-mark year, say, at the time of the census of 2001) should not be changed in this process.

Once the above courses of action are taken, there is no need to continue the scheme of the Urban Frame Survey (UFS) of the NSSO as this is just the duplication of work. The resources engaged in this activity may be utilised for carrying out usual surveys with increased sample size or for undertaking other activities like methodological studies.

While carrying out population censuses or economic censuses, a unique geo-coding structure for all the villages and EBs of the country (to be formed as per the guidelines given above) must be devised. That is, the same village or EB (both rural and urban) must be provided with the same code and it must never be changed. There are 6,35,109 villages in the country as per the population census, 1991. Thus an 8-digit code structure may be devised for codification of the villages and EBs in the rural areas. Of the 8 digits, the first 6 digits might be used for coding of villages and the last 2 digits for coding of EBs within each village. Another directory of codes (State/U.T. code x district code x tehsil code for each 8-digit code of village x EB) may be prepared to link the villages or EBs with the concerned State/U.T., district and tehsil to which the village or EB belongs. Similarly, a unique 6-digit code structure may be devised for the urban EBs required to be formed as per the suggestions given. As suggested for rural areas, another directory of codes (State/U.T. code x district code x town code x ward/charge/circle code) may be prepared to link each urban EB with the related State/U.T., district, town or ward/charge/circle. District x tehsil code (and not the village code or EB code) of any village/rural EB or district x town code (and not the EB code) of any urban EB may change in the subsequent censuses due to re-organisation of districts taking place from time to time. Due to growth of population over time, it might be necessary to split an EB into a number of EBs in the subsequent censuses. Necessary provision may be made in the code structure to link the erstwhile EB with the subsequently formed EBs.

The above suggestions may be implemented from the population census, 2001 itself, if possible. Once a unique geo-coding structure for all the villages / EBs is devised and implemented, a significant breakthrough in the existing system is achieved. With this, it becomes possible to judge the progress of different areas (say, villages / EBs, *tehsils*, towns, etc.) over time once the census or sample survey data relating to various characteristics at the desired geographical level are analysed for different time points.

**Interaction with Data Users**  A huge investment is made in collecting data through census and sample surveys. It is advisable that regular interaction is there between the data-collecting agencies and the data users regarding the usefulness/demand of such data and also on the periodicity or time interval at which the survey data on different subjects are required. If quality of the past available data on a subject is questioned by the users, proper methodology may be developed before collecting such data in the future.

**Standardisation of Concepts and Definitions**  Whenever more than one agency are involved in collecting data on the same subjects, it is necessary that they follow similar concepts and definitions so that the data thrown up by them are comparable. This will help in assessing the quality of field work of either agency at least when the data produced by them show different trends.

so that the data thrown up by them are comparable. This will help in assessing the quality of field work of either agency at least when the data produced by them show different trends.

**Strengthening of Training and Field Inspection** For proper implementation of instructions at the time of collection of data, the aspect of strengthening of training and field inspection need not be re-emphasised. In fact, this aspect assumes much more importance when the field work is carried out by deploying *ad hoc*/temporary resources.

**Re-allocation of Resources** Currently, for the surveys undertaken by the NSSO, the State/U.T.-wise sample sizes of first-stage units (i.e. villages and urban blocks combined) are determined in proportion to the number of investigators in the respective State/U.T.s (for any survey of one-year duration, actual sample size of a State or U.T. is generally 16 to 18 times the number of investigators allotted to the State or U.T.). Total sample size at the State/U.T. level is kept unchanged over the years as any change in the sample size would call for transfer of investigators from one State/U.T. to another in different rounds, thus causing various administrative or other difficulties. For technical justification of the existing method of State/U.T. level sample allocations, it appears that there is a need to re-allocate the total available resources (particularly, the Investigators and the Assistant Superintendents) of the Field Operations Division (FOD) of the NSSO among various States/Union Territories. As the NSSO generally undertakes household surveys and surveys on unorganised manufacture and trade, it is felt that the total available resources at the country level may be allocated among the States/U.T.s by considering their relative importance with respect to population and number of workers engaged in unorganised manufacturing and trading enterprises as per the latest data. There different allocations, viz. one in proportion to population and the other two in proportion to number of workers engaged in unorganised manufacture and trade, respectively, may first be obtained and then an ideal allocation may be worked out which might be the simple average of the three different allocations or some weighted average of them. While doing so, a minimum allocation must be ensured to each State/Union Territory.

**Quick Tabulation** Collection of data with a huge cost loses much of its relevance if the ultimate output, i.e. the survey results, are not made available within a reasonable time. Hence it is important to see that the size of the survey questionnaires is rationalised and no unnecessary items are included. Further, all necessary steps may be taken to validate the data and publish the results within a stipulated time period. In this context, it may be mentioned that there is a serious time lag in sofar as tabulation/report writing by many of the State Statistical Bureaus (based on the State sample data collected parallelly with the Central sample through the National Sample Survey) are concerned. Measures may be taken to improve this situation so that both the Central sample and State sample results for each State/U.T. are available within a reasonable time period after the survey. This achievement will help in pooling Central and State sample data for generating improved estimates or releasing estimates at further disaggregated levels, say, below the State/U.T. level.

**Coordination of Work** Last but not the least, it is essential that there is a regular interaction among all the data-collecting/disseminating agencies and the data users so far as the data coverage and standardisation of the concepts and definitions are concerned. This is required also for avoiding duplication of work and for a proper assessment of the demands for various types of data so that the scarce resources may be fruitfully utilised for collecting, processing and disseminating data which are in great demand.

# SOME CURRENT TRENDS IN SAMPLE SURVEY THEORY AND METHODS THAT MAY BE RELEVANT TO NSS

J.N.K. Rao[*]

The National Sample Survey Organization (NSSO) has been a leading sample survey organization since its creation in 1950 under the leadership of Professor P.C. Mahalanobis, who made pioneering contributions to the design of large-scale sample surveys and to the use of interpenetrating sub-samples for handling measurement errors. Survey statisticians formerly associated with the NSS, particularly D. B. Lahiri, M. N. Murthy, J. M. Sen Gupta and Des Raj, also made important contributions to the theory and practice of large-scale surveys. The NSSO continues to conduct major multi-subject surveys that provide valuable data required for planning and formulating policies.

Growing demand for reliable statistics at State and region levels has prompted the NSS to enlarge the sample size over the years by supplementing the "Central sample" with independent "State samples". In this context, two recent contributions may be relevant. First, dual frame methodology can be used to produce efficient estimates of totals and means by combining the Central and State samples optimally. Hartley (1974) developed a unified theory for dual frame surveys, but the sampling weights depend on the variable of interest. This implies a need to recompute weights for every variable of interest which is operationally inconvenient in practice when the survey involves a large number of variables, as in the case of surveys conducted by the NSSO. More importantly, such weights do not ensure consistency of figures when aggregated over variables, unlike a single set of weights computed and used for all variables. Recently, Skinner and Rao (1996) proposed a new method of estimation, based on pseudo-maximum likelihood, that uses a single set of weights and compares favourably in terms of efficiency to Hartley's and other methods that use variable-dependent weights. They also discussed variance estimation. Further results on variance estimation for dual frame surveys are given by Lohr and Rao (1997).

Secondly, M.P. Singh, also formerly associated with the NSS and currently Director of Household Survey Methods Division at Statistics Canada, has demonstrated how to exploit opportunities at the design stage to obtain significantly more efficient regional estimates at the expense of a small increase in the coefficient of variation at the provincial level (Singh *et al.*, 1994). He suggests reducing the number of sampling stages, using smaller strata and employing a compromise allocation to provide sufficient sample size at the region level. Adopting such strategies for the Canadian Labour Force Survey (LFS), he showed that the coefficient of variation (CV) for Unemployment Insurance regions can be significantly reduced at the expense of a small increase in the CV at the provincial level. The LFS used a monthly sample of 59,000 households and Singh's compromise allocation uses a core sample of 42,000 households to provide good estimates at the provincial level and the remaining sample of 17,000 households is allocated to produce the best possible sub-provincial estimates.

Even after using such compromise allocations and other measures suggested by M.P. Singh, the sample sizes in small areas (much smaller than a region) are rarely large enough for direct estimates to provide adequate precision. I understand that there is considerable interest in India to

[*] Professor of Statistics, Carleton University, Ottawa, Canada

estimate crop yields at the *tehsil* level (Tikkiwal, 1993). For reliable small-area estimation, it is necessary to borrow information from related small areas to find indirect estimates that increase the effective sample size and thus increase the precision. Such indirect estimates are based on either implicit or explicit models that link the areas through the use of supplementary data such as recent census counts and administrative records. An advantage of the explicit model-based approach is that it permits validation of assumed models and also provides reliable measures of variability associated with the indirect estimates. Ghosh and Rao (1994) provide an appraisal of small area methods based on indirect estimates and relevant references. Tikkiwal (1993) proposed a simulation method for producing small area statistics at the *tehsil* level. But this method does not borrow information from related *tehsils*, unlike the model-based methods. Also, one cannot increase the sample size (and thus decrease the variance) by simulation. In fact, simulation from a sample increases the variance!

Recent research has also focused on the efficient use of auxiliary information at the estimation stage. Auxiliary information in the form of known population totals (benchmarks) can be incorporated systematically through generalized regression estimates (Deville and Sarndal, 1992). More generally, we have calibration estimators that use weights which are as close as possible to the original weights, according to a specified distance measure, subject to benchmark constraints. We can now handle two or more post-stratifiers with known marginal counts through this approach. A drawback of calibrated weights is that range restrictions may not be satisfied; for example, there may be negative or extreme weights. Recently, Rao and Singh (1997) proposed a "ridge-shrinkage" method for range-restricted weight calibration.

It is often of considerable interest to estimate parameters more complex than the customary population or domain totals, means and ratios - for example, regression coefficients, median, poverty proportion (proportion below half the median) and logistic regression with a binary dependent variable. Such parameters can be formulated as solutions to "census" estimating equations, and then consistently estimated from the sample data using the design weights. Standard errors of such estimators from complex survey data can be readily estimated using a resampling method such as the jackknife method. Tests of hypotheses on the parameters can also be readily constructed, in particular using jackknife quasi-score tests (Rao, Scott and Skinner, 1998). Such methods take proper account of clustering, stratification and unequal probability sampling present in commonly used large-scale surveys, such as those conducted by the NSSO. It is also possible to use the Taylor linearization method, but software packages for survey data analysis, such as SUDAAN and PC CARP, cannot handle complex analyses with post-stratified weights, unlike the packages using the jackknife, for example WESVAR. Another advantage of the jackknife is that it uses a single standard error formula for all statistics, unlike the linearization method, which requires the derivation of a separate formula for each statistic. Other resampling methods, such as the bootstrap and balanced repeated replication (BRR), can also be used. These methods can handle non-smooth statistics (median, poverty proportion) unlike the jackknife, which is known to be inconsistent for such statistics at least for simple random sampling. I am currently conducting research on making valid inferences for the fractile graphical analysis proposed by Professor Mahalanobis, taking proper account of the survey design features, by employing the jackknife. Resampling methods are computer-intensive, but high-speed computing is becoming cheap and software packages are readily available; WESVAR can be down-loaded free-of-charge.

Researchers in subject matter areas often analyze survey data using standard software packages, such as SPSS and SAS, that ignore the design features in calculating standard errors and test statistics. The NSSO should make an effort to educate these researchers in using proper methods for analysing the data collected by the NSS from a variety of surveys.

I have focused so far on sampling variability, but the importance of measurement errors and other non-sampling errors should be recognized in conducting any large-scale survey. Professor Mahalanobis, as noted before, used interpenetrating samples to handle measurement errors, and the NSS continues to pay attention to non-sampling errors in the design of surveys. Linacre and Trewin (1993) address the issue of resource allocation in the context of total survey design. They demonstrate that by optimizing resource allocations so that resources are deployed where error reduction is most effective, superior survey designs can be obtained. The NSS can benefit by adopting similar methods.

Past research on measurement errors focused on simple estimates, such as the sample mean, using additive models. Under this set-up, the estimates remain unbiased and the total variance can be estimated, e.g. using interpenetrating sub-samples. However, the estimates are no longer unbiased in the case of distribution functions and regression equations which are of considerable practical interest (Fuller, 1995). Fuller shows that by taking replicated observations on a sub-sample of individuals, the bias can be eliminated. Thus, both replication and interpenetration are needed to handle measurement errors in estimating complex parameters.

## References

Deville, J. C. and Sarndal, C. E. (1992). "Calibration estimators in survey sampling", *Journal of the American Statistical Association, 87*, 376-382.

Fuller, W. A. (1995). "Estimation in the presence of measurement errors", *International Statistical Review, 63*, 121-147.

Ghosh, M. and Rao, J.N.K. (1994). "Small area estimation: an appraisal", *Statistical Science, 9*, 55-93.

Hartley, H.O. (1974). "Multiple frame methodology and selected applications", *Sankhyā, Ser. C, 36*, 96-118.

Linacre, S.J. and Trewin, D.J. (1993). "Total survey design - application to a collection of the construction industry", *Journal of Official Statistics, 1993*, 611-621.

Lohr, S.L. and Rao, J.N.K. (1997). "Jackknife variance estimation in dual frame surveys", *Proc. Survey Research Sec.*, American Statistical Association (in press).

Rao, J.N.K. and Singh, A.C. (1997). "A ridge-shrinkage method for range-restricted weight calibration in survey sampling", *Proc. Survey Research Sec.*, American Statistical Association (in press).

Rao, J.N.K., Scott, A.J. and Skinner, C.J. (1998). "Quasi-score tests with survey data", *Statistica Sinica* (in press).

Singh, M.P., Gambino, J. and Mantel, H.J. (1994). "Issues and strategies for small area data", *Survey Methodology, 20*, 3-32.

Skinner, C.J. and Rao, J.N.K. (1996). "Estimation in dual frame surveys with complex designs", *Journal of the American Statistical Association, 91*, 349-356/

Tikkiwal, B.D. (1993). "Modelling through survey data for small domains", *In Small Area Statistics and Survey Designs*. Central Statistical Office, Warsaw, pp. 129-155.

# CREATION OF GEOGRAPHICAL INFORMATION DATABASE USING POSTAL INDEX NUMBER

## A. J. Roy[*]

## Need

Planners and administrators require to use geographical database in order to take decisions based on regional variations for allocation of funds and to monitor various parameters of interest in public administration. Further, presentation of geographical database with representative points on maps and charts provides a very useful tool in the hands of a decision-maker.

## Approach

The Indian Postal Department has adopted the PIN Code for identification of specific areas under the lead post office for providing service under them. This provides us with an interesting opportunity to identify each area in a unique way without confusion over names so that the lead post office can be used as a reference point on the geographical area under its coverage for a given PIN code. Thus, by maintaining a database of all the representative post offices with longitude and latitude positions, we can uniquely identify geographical reference points and their corresponding coverage areas using the PIN code. So, by including the PIN code with any geographical information, we can augment geographical database very easily. The utility of such a database will be enhanced if we maintain it on public computer network.

## Application

We may develop useful databases augmented with the PIN codes by collecting data mainly from the many agencies and apparatus in the government such as local bodies to Central Government ministries. The important subjects broadly to be covered are as follows:

1. Law and Order Administration
2. Human Resources
3. Health & Family Welfare
4. Power & Energy Consumption
5. Environment and Forests
6. Industry & Mining
7. Banking & Finance
8. Transport

[ Some useful examples of such databases are included in the Appendix-II ]

## Implementation

The creation of such a vast database at the national level as such would require initiatives to be taken by different agencies working at the national level like those under Central Government ministries. Also, the cost factor can be prohibitive if not planned properly during the early stages of implementation of such a database. If we focus on those information and statistics which are normally produced as a by-product of the routine administrative activities, then we may build up

[*]Deputy Director, DPD, NSSO, Calcutta. Views expressed in this article are those of the author and not necessarily of the organisation to which he belongs.

useful databases within acceptable cost limits. Subsequently, a part of such costs can be recovered by identifying salable data for paying users.

Initially, Central ministries may be approached for financial assistance in the creation of such databases on computer systems linked electronically over digital networks such as NICNET, etc. Some examples of thrust areas for different Central ministries are stated. The Finance Ministry may provide funds for creation of database at each PIN code level with respect to loan and credit repayment through banks. Similarly, the Home Ministry may monitor crime incidence for each area and detect activities of criminal groups moving from one area to another. The Health Ministry may identify areas not adequately provided with medical services and facilities by the local administration besides pinpointing and monitoring outbreak of infectious and contagious diseases. Also, the Ministry of Environment and Forestry may create area-wise land-use database as reported by local officials and update such data with satellite imageries obtained from the Indian Space Research Organisation.

To start with, the population census of India can be used to build up such databases by providing area-wise information on important parameters so that all important statistics from the census operation are available at each PIN code level. Also, each area identified and listed in the process should be available to the public with the corresponding unique PIN code. Hence, this database from the census operation may be used as a users' reference by other data collecting agencies.

### Limitation

PIN codes were created with the objective of improving the efficiency of Postal operations and, therefore, have their limitations when their use is extended to other purposes. Through the PIN code we can identify a major administrative area such as a district from the first 3 digits and by using all the 6 digits we can identify a specific postal service area. However, the size of such postal service area is said to depend on the quantum of workload for that area so that no service area has to deal with an excessive load.

Thus, it may be suggested that the number of such postal areas in a region indirectly corresponds to total activities in that region. For example, urban areas have higher concentration of PIN code areas than rural areas of the same geographical size. [See Appendix-I]

It is also desirable that the administrative area up to a *tehsil* or its equivalent should be provided with a unique PIN code. If necessary, we may extend a PIN code attaching a unique letter of the alphabet to the right to represent a specific area. This way we may specify smaller areas with extended PIN codes. Also, careful efforts should be made to clearly demarcate each PIN code area with its associated public administration units avoiding overlaps to guard against possible confusions.

### Conclusion

Hopefully, the next census operation around the year 2000-1 may be used as the first stepping-stone for effective use of PIN codes in the creation of geographical database. This historical opportunity should be exploited to the full to gain acceptance of PIN codes for wider use by other data collecting agencies like the National Sample Survey Organisation, MARG etc. Such a strategy could also be adopted by other national census agencies too as such a use of PIN code provides a cost-effective option to create geographical information database across the country.

### References

1. *All India Directory of Postal Index Number, Department of Posts*, Govt. of India, 1992.
2. *National Atlas of India*, National Atlas and Thematic Mapping Organisation, Deptt. of Science & Technology, Govt. of India.

**Appendix-I**

**PIN Code: An Outline**

The Postal Index Number Scheme, also known as the PIN Code in short, was introduced by the Post & Telegraphs Department on 15 August, 1972. Under the scheme, every Head Post Office and Sub-Post Office which delivers mail has been allotted an individual six-digit number. The Branch Post Offices use the numbers allotted to their parent offices. In our country having diverse languages and scripts, the introduction of the digital code was supposed to ease the problem of sorting mail by sorters. Further, there are several different towns in this country bearing the same name. So the addition of the PIN code after the address enables the sorter to identify the destination correctly, thereby reducing chances of misdirection and delay in postal deliveries.

The PIN code digits, when read successively from the left to the right, pin-point and locate the post office geographically. The country has been divided into eight different zones. The first digit indicates the zone. The first three taken together indicate the sorting unit under which the office falls. The last three specify the particular delivery Post office under the sorting unit.

Such coding systems are adopted by many other countries as well e.g. ZIP code by the US Postal Service, National Postal Code by the Postal Service of Canada, etc.

**Appendix-II**

**Examples of Useful Databases**

The following list of information may be obtained areawise for the creation of geographical database from different agencies in operation:

1. Statistics on industrial production by input and output (special reference to the Annual Survey of Industries data)
2. Surveys on origin and destination of vehicles as used in the transport sector
3. Information on yield of crops
4. Information on surveillance of pest/disease vector control and management from health authorities
5. Meteorological data such as rainfall, temperature, humidity from the Meteorological Department
6. List of tourist destinations along with infrastructural facilities to be used by tourists from Tourism Department/Ministry
7. List of beneficiaries under various government programmes providing subsidies from Department/Ministry of Planning
8. Information on availability and consumption of various services provided by different utilities - electricity, telecommunication, etc. from commission/department controlling such utilities
9. Information on intensity of use of pesticides/inorganic fertilizers with special reference to subsidies from Ministry of Agriculture/Petrochemicals
10. Information on availability of primary schools/post-primary schools/colleges & higher education centres from Education Department/Ministry
11. Information on availability of water for drinking by source and by quality with special reference to arsenic and fluoride content from health authorities
12. Information on health care facilities to population control measures, etc from health authorities
13. Information on crime, incidence, tracking criminals, fugitives and stolen vehicles, etc from Home Department/Ministry
14. Information of surveillance on outbreak of communicable diseases from health authorities

# STATISTICS AND INFORMATION
# AN ANALYSIS OF THE PRESENT AND PERSPECTIVES FOR THE FUTURE

Ashish Kumar[1] and T. R. Sreenivas[2]

*" The government are very keen on amassing statistics. They collect them, add them, raise them to the nth power, take the cube root and prepare wonderful diagrams. But you must never forget that every one of these figures comes in the first instance from the village watchman who puts down what he damn pleases. "*

Perhaps many statisticians in India also share the cynicism expressed in the above words by Sir Josiah Stump in some other context. In the golden jubilee year of Indian Independence there is a pressing need even among the foremost of Indian statisticians to look at the whole statistical system in India afresh and try to acknowledge the contributions made by the statisticians of the yore to the community in general and the public policy and planning in particular. This is in the backdrop of fears expressed in many quarters that Statistics, albeit having a very short history of being a separate academic discipline, may soon vanish as a separate subject to be taught at undergraduate and post-graduate courses. An attempt is made here to analyse the present status and understand the possible future scenario of the statistical system as a whole.

## 1. Statistical System

A statistical system can be broadly identified as consisting of the following sub-systems:

1. Education and training.
2. Central statistical agencies - like the Central Statistical Organisation, the National Sample Survey Organisation, the Office of the Registrar General.
3. State statistical agencies - like a State Statistical Bureau and its associate bodies in various departments in the State governments.
4. Central agencies attached to various functional organisations - like statistical offices in Central ministries, the Department of Statistical Analysis and Computer Services (DESACS) in the RBI, Textile Committee, statistical division of the NCERT.
5. Private agencies providing statistical information - like the CMIE, ORG, IMRB.
6. Autonomous or quasi-governmental agencies providing statistical information - like the National Council of Applied Economic Research (NCAER), National Institute of Rural Development.

Further, there are many other organisations mainly involved in providing user-specific data. With a plethora of organisations involved in data collection and dissemination, one can really say that India has a wealth of data and is adding to its wealth regularly. It is indeed a tragedy that the efforts of all these organisations are resulting in providing data and more data but not adequate information. Data are getting generated by routine administrative systems with little appreciation of

---

[1] Joint Director, DPC, NSSO, Bangalore
[2] Deputy Director, DPC, NSSO, Bangalore
Views expressed in the article are those of the authors and not necessarily of the organisation to which they belong.

the importance of the data collected. The shadow of the village watchman is perhaps looming large over all these efforts.

## 2. Data and Information

It is elementary to distinguish data from information; information is distilled data and the processes of distillation include classification, tabulation, graphical or diagrammatic representation, inference; in a nutshell, most of modern statistical technology applied to data yields information. It is an acknowledged fact that information has become a key resource in modern-day decision-making. Computers have greatly facilitated the conversion of data into information; they have revolutionised the field of information in the same way as they have revolutionised every other walk of human life with their speed and number-crunching ability. But for the developments in the computer field, many of the statistical techniques like regression analysis, reliability calculations or, for that matter, the now so popular techniques of Initial Data Examination would have remained mainly academic pursuits. But it is naïve to equate computers with information. The power of information lies in its inherent features; they can be summarised as follows:

1. Information needs accuracy, timeliness and perspective. These attributes of information are dependent on the need for information. Higher-than-needed accuracy and longer-than-needed perspective result in wastage of resources and cause information overload and fatigue. Lesser in these attributes, information may be plainly useless in a given context. It is always useful to have current information but even here there are tolerance limits imposed by the need of the user.

2. Information can be broadly classified as strategic, management and operational information. Strategic information is required by planners and policy-makers to set objectives, mission statements and to frame policies. Information needed in this context is on broad and relevant areas. But the time horizon of information is vast. It may span many years. However, there are broader tolerance limits on time frame. Many a time forecasts based on a-year-old data are acceptable whereas in some other situations such forecasts may not be reliable even if based on a fairly long series of data. The Planning Commission uses this kind of macro-economic information for national-level plan formulation.

3. Management information is required for monitoring and surveillance. The information needed is at a much finer level of aggregation but with a smaller perspective. There will be more stringent limits on timeliness. Perhaps monthly data or sometimes even weekly/quarterly data are needed for this purpose. Many on-going infrastructure /developmental projects use this kind of information.

4. Operational information is needed by the person who is on the scene of action. Even a statistical investigator collecting data or data entry operator keying in hundreds of records per day should have help of a mechanism which provides him continuous feedback on accuracy and progress of his work. The kind of information needed is minimal in aggregation but high in accuracy. And it is required with almost no time lag.

It is impossible that a single centralised organisation caters to all the information needs. At the same time, it is highly desirable that the information provided by various agencies conform to some specified standards. There should be clearly defined rôles for various persons/organisations involved in the task of providing information. It is obvious that the level of functional knowledge, maturity and understanding required by the person who is providing information varies and depends on the level for which the information is being provided. One providing strategic information must be a real visionary who has a good grounding in the subject area of information and who has a reputation and position for decision-makers to accept his word; in fact, he must be able to operate as

a member of the decision-making team as an expert in the field of information. The levels of personnel required for other informational assignments also would vary accordingly.

## 3. Rôle of Education and Training

A clear conception of a statistician as an expert in information provides for an ideal background for evaluation of the present education and training to budding statisticians in India. Most of the post-graduate programmes in statistics have not been extremely successful to attract the best of talents. Further trained statisticians have been lost to other emerging disciplines like computer science, econometrics, etc., since statistics as a profession could not promise any better prospects for growth and development. The uncompromising obsession with mathematics or mathematical statistics as the corner-stone for teaching all statistics courses have deprived students of statistics of all other necessary skills for their survival. Even persons of repute like Dr. C.R.Rao admit that a good knowledge in at least one other functional area like economics is essential for a student of statistics to test and polish his skills. But there are very few universities which have devised their curriculum on these lines. In fact, many universities are still reluctant to expose statistics students to computers in order not to dilute their courses! It is also essential that students of statistics are exposed thoroughly to the philosophy of knowledge - epistemology - so that they appreciate the limitations of empiricism and deductive logic, for which the technology of statistics provides the necessary tools.

Further, students of statistics at the post-graduate level need to specialise in any one subject - there should be at least two full papers on a chosen subject like economics, public health, agriculture, mathematics, marketing – to be taught by experts in the subject in their own settings. A course-ending project work in the chosen subject in real-life setting should be part of the curriculum. This kind of a course will not only do justice to proper teaching of the subject but would produce better-equipped professionals.

Another disturbing trend which has been observed in higher education has been that professionals from functional areas have started teaching statistics to their students. This may be due either to the inability of the average statistics teacher to relate his knowledge to the students of the particular professional course or to the reluctance on the part of others to accept statistics as a separate discipline warranting specialised coaching, or to both. In any case, it is not an effective way of teaching the subject and should be done away with if students of other professions would like to have some skills in statistics.

In a similar fashion, training imparted to statistical personnel should be task and subject oriented. Involvement of statistical personnel as information experts will definitely provide much-needed focus to many endeavours which would result in optimal utilisation of scarce resources.

## 4. Indian Statistical Service (ISS)

Since it is a Central "Group A" service, training in the ISS should be a rôle model for training imparted to statistical personnel elsewhere. Although as a multi-departmental service, its officers are posted in various ministries, resulting in the need for differing functional expertise, the service has not developed suitable training programmes to meet such requirements. There is a need to develop a core curriculum based on skill essential to offices where such officers are posted. They should be allowed to specialise in some functional areas. The training programme should culminate with a real-life project work. Such a training will be able to equip them to understand problems in their chosen functional area and enable them to provide suitable statistical and information-based solutions.

In short, the training should aim at creating officers who can provide information leadership in various ministries. In order to facilitate such a rôle for the ISS officers, job content and job design may have to be suitably modified and enriched. It may also be useful to look afresh at the recruitment policy for ISS officers with a view to attracting personnel with an aptitude for learning other disciplines. It may be a good idea to select officers in the service through the Combined Civil Services examination with a qualification that the person should have taken statistics as one of the optional subjects.

## 5. Rôle Definition of Government Statistical Agencies

Pondering over the functioning of various Central and State statistical agencies, one can easily stumble upon the lack of clear rôle definition and the lack of effective co-ordination as the two most worrisome aspects of the system. It is often the case that one Central agency does not accept the data of another. Duplication of work is almost accepted as inevitable and the results of different surveys on the same subject are mostly non-comparable, not only due to different settings of time and place but also due to conceptual ambiguities, design limitations and interpretation fallacies. Looking from the angle of information, it is prudent that the rôle of various agencies in providing different levels of information may be clearly and effectively defined.

It would be ideal if dedicated Central/State statistical agencies were made fully responsible to provide the strategic information. Statistical agencies in ministries should be responsible for management information and individual executing agencies should be responsible for operational information. A representative of the Central statistical agency present in each functional ministry can play the co-ordinating rôle between the ministry and the Central agency on the one hand and guide suitable information collection system within the ministry. This would reduce, if not remove, conceptual ambiguities, rôle definition problems, etc.

## 6. Statistics as a Profession

In the context of differing agencies providing a plethora of information - often not consistent with each other - it is suggested that the Department of Statistics, Ministry of Planning & Programme Implementation, Government of India assert its assigned rôle of laying down and maintaining norms and standards involving concepts, definitions and methodology in the collection of data and dissemination of information. It is desirable to create a professional body similar to professional institutions for accountants, company secretaries, etc., and try to create a manpower of qualified statisticians who may be called chartered statisticians. It has to be made mandatory that all projects involving governmental funding be cleared by these professionals as conforming to specified standards of data collection and monitoring.

Perhaps the much desired system of Statistical Advisors in each Ministry of Central Government and in each State reporting to the Central Statistical Commissioner in Central Government, when it comes into existence, will be able to take care of such requirements whenever the projects are executed by Government agencies themselves.

Necessary amendments to relevant statutes e.g. the Collection of Statistics Act, may have to be taken up on a priority basis if wastage of resources for collection of information is to be avoided. Statistical professionals can play an important rôle in implementation of the government's will to provide right to information for every citizen.

## 7. Status of Technology Use in Statistical System

Information technology and communication technologies have greatly advanced. These advancements have been very well adapted to the statistical and information analysis systems in the

private sector, multinational organisations, etc. These technologies, though an integral part of information management, have not percolated to the government and semi-government organisations. The backwardness of the statistical sector in the adoption of the modern electronic technologies has resulted in delays and ineffective analysis systems. Considering the vastness of the scope and the massive effort required by the statistical organisations in the government, there is a need to urgently modernise the sector with the latest computer systems along with communication systems. In fact, there is a need to have a dedicated satellite-based communication network for the statistical information system. There is also a need to modernise data collection systems where field investigators could be equipped with tape recorders, video cameras and global positioning systems. Similarly, for data and information dissemination, there is a need to adopt better media, including CD-ROMS, maintenance of Websites, use of Internet, etc. With all this modernisation, the statistical organisations should start using advanced statistical and information analysis methods to provide clearer and deeper understanding of the phenomena being reported upon.

## 8. Conclusion

We have attempted an evaluation of the present statistical system in the country in the light of present-day technology and emerging demands. It is in no way an attempt to belittle the contributions made by statisticians and information experts in the past. Our effort is only to make suggestions so that one of the most backward sectors, in spite of its inherent importance, does not continue to remain backward and some steps are taken to revitalise it.

# THE ANNUAL SURVEY OF INDUSTRIES – SOME REFLECTIONS

G.S.Lakshmi[1], Sewa Nand[2] and M.S.Maulik[3]

## 1. Brief History

The Industrial Statistics Act, 1942 empowered the Government to collect data from factories defined under the Factories Act, 1934, a factory being an establishment employing 20 or more workers and using power. A Directorate of Industrial Statistics was set up under the then Ministry of Commerce and Industry in 1946 to coordinate the work, to advise on technical and legal matters and to process and publish the data on an all-India basis. The Directorate was transferred to the CSO in July 1957 and came to be known as the Industrial Statistics Wing.

The Factories Act, 1934 was later revised and replaced by the Indian Factories Act, 1948. Under the new act, the definition of a factory was extended to cover establishments with 10 or more workers and using power and 20 or more workers but not using power. The Industrial Statistics Act, 1942 was replaced by the more wide-ranging Collection of Statistics Act of 1953. Industrial statistics from the factories registered under the Indian Factories Act, 1948 are now being collected under the powers conferred by the Collection of Statistics Act, 1953.

The Directorate of Industrial Statistics launched the Census of Manufacturing Industries (CMI) in 1946 with the object of studying the structure of the Indian industry and estimating its contribution to national income. Because of practical difficulties, the CMI could cover only 29 of the 63 industry groups specified in the Industrial Statistics Act and extended only to 11 States of the Indian Union. By 1958, the regional coverage of the CMI extended to 13 States and two Union Territories.

Following the recommendation of the National Income Committee (1949), the Directorate of Industrial Statistics conducted the first Sample Survey of Manufacturing Industries (SSMI) in 1949 for collecting data from factories falling under the 34 industry groups left out by the CMI and defined under the Factories Act, 1934. The technical work, including the survey design, sample selection and preparation of schedules, was undertaken by the Directorate while the tabulation and analysis of data was carried out by the Indian Statistical Institute. Similar surveys were conducted annually until 1958 by the then Directorate of National Sample Survey.

## 2. Description of the ASI

(a) **Introduction** The Collection of Statistics (Central) Rules, 1959 framed under the 1953 Act provided, among others, for a comprehensive Annual Survey of Industries (ASI) in India. This survey replaced both the CMI and the SSMI. The ASI was launched in 1960 with 1959 as the reference year and is continuing since then except for the ASI 1972. For the ASI the Collection of Statistics Act, 1953 and the Rules framed thereunder in 1959 provide the statutory basis. The ASI refers to the factories defined in accordance with the Indian Factories Act, 1948 and thus has a coverage wider than that of the CMI and the SSMI put together.

[1] Deputy Director, Industrial Statistics Division, CSO, New Delhi
[2] Joint Director, Industrial Statistics Division, CSO, New Delhi
[3] Head, Industrial Statistics Division, CSO, New Delhi
Views expressed in the paper are those of the authors and not of the organisation to which they belong.

(b) **Scope and Coverage**  Geographically, the ASI covers the entire country except some North-Eastern States (Arunachal Pradesh, Mizoram and Sikkim) and the Union Territory (U.T.) of Lakshadweep. The frames used for the survey are as under:

(i) all factories registered under Sections 2m(i) and 2m(ii) of the Factories Act, 1948 employing 10 or more workers and using power as well as those employing 20 or more workers but not using power;

(ii) *bidi* and cigar manufacturing establishments registered under the *Bidi* and Cigar Workers (Conditions of Employment) Act, 1966 with coverage of units as in (i) above;

(iii) electricity undertakings engaged in generation, transmission and distribution of electricity registered with the Central Electricity Authority irrespective of their employment size; and

(iv) certain service units like water supply undertakings and cold storage units.

(c) **Updation of Frame**  The frames are maintained by the Chief Inspectors of Factories (CIFs) for the units registered under the Factories Act, 1948 or by licensing authorities in case of *bidi* and cigar establishments and electricity undertakings. The regional offices of the FOD (NSSO) maintain close liaison with the CIFs in updating the frame every year, which is finally revised every three years. The frame was last revised in 1995-96.

(d) **Sampling Design**  The ASI is based on a mixed design of complete enumeration (census) and sampling of units as explained below :

(i) all factories (units) situated in industrially less developed States/U.T.s (viz. Goa, Himachal Pradesh, Jammu and Kashmir, Manipur, Meghalaya, Nagaland, Tripura, A & N Islands, Chandigarh, Dadra & Nagar Haveli, Daman & Diu and Pondicherry) are covered on census basis;

(ii) all electricity undertakings are covered on census basis;

(iii) all factories employing 100 and more workers are covered on census basis;

(iv) in the case of factories with 10-99 workers, if the number of factories is less than 20 at 3-digit level of industry in a State/U.T., then the 3-digited group is covered on census basis; and

(v) the remaining factories are surveyed on sampling basis. These factories are stratified into State X 3-digited level of industry and three subgroups are considered;

(a) subgroup with less than 20 factories in which case census is adopted as already mentioned in (iv) above;

(b) subgroup with 21-60 factories from which a fixed sample of 20 units is chosen by adopting circular systematic sampling;

(c) subgroup with 61 or more factories from which a fixed sample ratio of one-third is selected by adopting circular systematic sampling.

Thus, by rotation over a period of three years all the factories are surveyed and the frame is revised after the three-year period.

(e) **Reference Period**  In the beginning, the reference period in the ASI was the calendar year for all items except fixed assets and working capital for which it was the accounting year. After a careful consideration of all the relevant issues, the reference year was changed to the financial year (April – March) for all items since the ASI of 1988-89.

(f) **Industrial Classification**  The classification adopted for the ASI was the Standard Industrial Classification from 1963 till 1973-74. With effect from ASI 1973-74, the National Industrial Classification (NIC) 1970 was adopted. The NIC 1987 is being followed from the ASI of 1989-90.

(g) **Data Collection**    The Director, FOD (NSSO) has been designated as the Statistics Authority under the Collection of Statistics Act, 1953. The field work is carried out by the FOD (NSSO) through its network of regional and subregional offices located in various parts of the country. Notices are issued by the FOD (NSSO) to owners of the factories enclosing a complete set of the schedule and instructions requiring them to submit the returns pertaining to the previous financial year by a specified date. The data collection itself is spread over from September/October and spills over to the next June/July and returns are dispatched to the tabulating agencies after necessary checks at the field offices of the FOD (NSSO).

Data under the ASI are collected through a comprehensive schedule. In the initial rounds, the schedule sought particulars relating to manufacturing activity only. Over the years, additions were made to meet the specific data requirements of various organisations. By 1973-74, the schedule consisted of five parts: Part I on manufacturing, Part II on labour turn-over, Part III on stocks and consumption of components and accessories in small scale sector, Part IV on constructioon expenditure, and Part V on indirect taxes, sales subsidies and capacity of power equipment installed. It was felt that the ASI schedule had become too unwieldy and complicated. So a modified schedule with three parts on manufacturing, labour and construction was adopted with effect from ASI, 1974-75. The schedule was last modified in 1987-88.

(h) **Tabulation and Dissemination of Data**    The ASI results are tabulated as 'summary' and 'detailed' results through scrutiny of error lists and trial tables generated through computers (PCs).

Up to the ASI, 1994-95, summary results were processed in the Computer Centre, New Delhi and the detailed results in the Industrial Statistics Wing(IS Wing) of the CSO at Calcutta. From ASI 1995-96 onwards, both summary as well as detailed tabulation is being undertaken in CSO (IS Wing), Calcutta.

The summary results are published every year in two volumes with a gap of 10 to 12 months from the date of receipt of the last batch of returns disseminating important parameters at two- and three-digit level of industry at all-India and States. In disseminating the results, secrecy is maintained about individual units as stipulated in the Collection of Statistics Act, 1953. Thus, certain merger plans are followed in cases where the number of units in any particular industry is less than three.

Till 1989-90, the detailed results at the 4-digit level of industry were processed quinquennially; from 1993-94 onwards, the detailed results are being processed and disseminated at the 4-digit level of industry for various States /U.T.s annually. The detailed results are released in 16 volumes.

(i) **Some salient features**

Some important technical coefficients for factories by year of initial production are given below :

| Year of initial production | Fixed capital to NVA | Fixed capital to output | NVA to output | Emoluments to NVA |
| --- | --- | --- | --- | --- |
| Up to 1969 | 2.68 | 0.64 | 0.24 | 0.42 |
| 1970 - 1979 | 1.63 | 0.30 | 0.19 | 0.33 |
| 1980 - 1989 | 2.21 | 0.43 | 0.19 | 0.21 |
| 1990 - 1999 | 4.47 | 0.80 | 0.18 | 0.19 |

It may be seen that the fixed capital to net value added ratio, which provides a measure of the capital required to produce one unit of net output (i.e. net value added), was a maximum (4.47) for the factories which started their production in the current decade whereas the ratio was a minimum (1.63) for the factories which started their production during 1970-1979. It appears that the current industries are consuming more and more fixed capital with no extra contribution in net value added. This can be studied in more detail for particular industry and particular region. The table also shows that the share of labour in net value added measured by the ratio of emoluments to net value added varied from the lowest 0.19 for the factories which started their production in 1990-1999 to 0.42 for the factories which started their production in the initial years (up to 1969).

Some important technical coefficients for the years 1980-81 to 1994-95 are presented in the table below :

| Year | FC to NVA | FC to Output | NVA to Output | Emoluments to NVA |
|---|---|---|---|---|
| 1980 -1981 | 2.51 | 0.49 | 0.20 | 0.51 |
| 1981 - 1982 | 2.39 | 0.47 | 0.20 | 0.47 |
| 1982 - 1983 | 2.46 | 0.48 | 0.19 | 0.48 |
| 1983 - 1984 | 2.41 | 0.51 | 0.22 | 0.45 |
| 1984 - 1985 | 2.63 | 0.52 | 0.20 | 0.51 |
| 1985 - 1986 | 2.66 | 0.50 | 0.19 | 0.49 |
| 1986 - 1987 | 2.63 | 0.51 | 0.19 | 0.48 |
| 1987 - 1988 | 2.77 | 0.51 | 0.18 | 0.50 |
| 1988 - 1989 | 2.57 | 0.48 | 0.19 | 0.45 |
| 1989 - 1990 | 2.51 | 0.46 | 0.19 | 0.43 |
| 1990 - 1991 | 2.59 | 0.49 | 0.19 | 0.40 |
| 1991 - 1992 | 2.77 | 0.51 | 0.18 | 0.38 |
| 1992 - 1993 | 2.71 | 0.52 | 0.19 | 0.39 |
| 1993 - 1994 | 2.54 | 0.53 | 0.21 | 0.32 |
| 1994 - 1995 | 2.56 | 0.54 | 0.21 | 0.33 |

It may be seen that the capital-output ratio, which provides a measure of the capital required to produce one unit of net output (i.e. net value added), varied from 2.39 in 1981-82 to 2.77 in 1987-88 and 1991-92. The capital required to produce one unit of gross output had gone up to 0.54 in 1994-95 from 0.47 in 1981-82. The level of efficiency measured by the ratio of the net value added to gross output varied from 0.18 in 1987-88 to 0.22 in 1983-84. The share of labour in net value added, measured by the ratio of emoluments to the net value added, varied from 0.32 in 1993-94 to 0.51 in 1980-81 and 1984-85.

## 3. Proposals for Improvement

To reduce the time lag in releasing the ASI results and enhancing their accuracy, it has been decided that from the ASI, 1995-96 onwards, the entire processing of summarising as well as detailed tabulation of ASI data would be taken up by the CSO (IS Wing), Calcutta. Up to the ASI of 1994-95, summary tabulation was done in the Computer Centre, Deptt. of Statistics, New Delhi.

A symposium on the revision of sampling design for the ASI was held in Calcutta in March 1998. At the symposium, it was decided that the work of revision of sampling design and reduction in size of the ASI schedule should be adopted for ASI, 1997-98. Also, a monthly survey of large and important establishments is being thought of. Follow-up action to reduce time lag in data collection /processing as well as introduction of monthly survey is being taken up.

The field collection of data is being streamlined from 1997-98 onwards in such a way that the data collection time is to be reduced drastically from more than 15 months as in the past to only 60 days. In particular for 1997-98, it has been decided that the entire data collection with regard to the ASI work will be over during October-November, 1998 and all returns are expected to be received by the IS Wing, Calcutta by the middle of December, 1998.

Improvement in quality and enforcement of a strict time schedule is also planned by the DoS by forming "Data Improvement Teams" from amongst senior officers of the Department who would undertake field visits to enforce strict discipline in regard to completion time for canvassing the schedules as well as for checking the quality of data collected.

# LARGE-SCALE SOCIO-ECONOMIC SAMPLE SURVEYS SOME METHODOLOGICAL ISSUES

S. P. Sharma[*]

### 1. Introduction and Background

Sample Surveys are being used for collection of primary data in the field of agriculture, industry and other socio-economic subjects like health, consumer expenditure, prices, debt, investment, etc., over the years. Sampling technique being a statistical method for estimating the behavioural aspects of various entities was in the knowledge of a limited number of professionals in the past. As such, only economists, statisticians and research workers were mostly the traditional users of the data thrown up by sample surveys. However, with the increasing level of education, growing complexities of administration, management of public sector as well as private sector enterprises and particularly after the availability of faster computing equipment, the use of sample surveys has grown in almost all areas and the large cross-section of society is using the results of sample surveys for various purposes. With the growth in population and development, increasing urbanisation, industrial development and specialisation in marketing, advertisement, communication and other activities, the rôle of information has become much wider. The timeliness and the quality of information are of the utmost importance.

It is truly called the information age with the state-of-the-art technologies at hand in the field of computing as well as information dissemination, with the help of inter-net, e-mail and multimedia facilities at our doorstep. While timely and quality information has a tremendous market, the same if delayed or with suspect quality may create a very bad impression about the agency supplying the same. It is, therefore, to be emphasised that timeliness and quality of information is to be ensured at any cost, while planning for any sample surveys.

### 2. Dealing with Size and Heterogeneity

With the growth in size of any population either that of industries, farms, households, enterprises or institutions in the country like India with its vast geographical area, natural physical and environmental diversity with economic and climatic extremes present in the conditions, the sampling exercise becomes complex. As such, to ensure the quality of results, the importance of adequate stratification and appropriate sample size need not be over-emphasised. With the right amount of stratification and with the help of adequate prior knowledge of the population to be surveyed based on previous experience, the over-all sample size can be reduced without sacrificing the quality of the estimates. It is, therefore, very important to stratify the population up to the adequate level on the objective basis, based on the criteria so as to minimise the intra-stratum variance and to maximise the inter-stratum variance. While deciding the sample size in different strata, adequate emphasis ought to be placed on the average cost of collection of data per FSU/SSU and sampling variance within the stratum rather than on the number of investigators available in the concerned area or on the proportion to the related population of the concerned entity in that region. Since the estimated population in different zones/regions/ sub-regions may vary from round to round

[*] Joint Director, DPC, NSSO, Ahmedabad. The views expressed in the paper are of the author and not of the organisation to which he belongs.

depending on the problem under consideration and the distribution of sample size may vary significantly from round to round across the zones/regions, there may be a necessity of re-allocating the number of investigators and other supervisory staff in different zones/regions in the Field Operation Division (FOD) based on population, industrialisation and intra-region heterogeneity in development/infrastructure. The sample size in zone should not depend on the number of investigators and other staff; rather, it should be the other way around. Before deciding about the sample FSUs/SSUs over the various zones/regions, the detailed study of the population characteristics may be made with the help of the previous records of the same and the heterogeneity in the principal characteristics concerned in the different zones particularly considering the topic under study in a given round.

## 3. Minimum Sample Size

For any problem under study, the minimum sample size for every zone/region may be decided in such a way that it should become significant at the State level at least, if not at the sub-division level. Since, with the decentralisation of planning, while more emphasis is being laid on regional issues for proper execution of various schemes, it is necessary to provide information at the State/divisional level. Further, while over-all results at the national level may be appropriate for allocating funds at the national level, the distribution of the same across and within the States cannot be done properly in the absence of appropriate estimates at that level. The diversity within States and between rural and urban areas is also very significant and is growing with the development particularly after liberalisation of industrial and trade policies since 1991. So it is also necessary to fix different techniques and proportions for sampling methodology and sample size in different States/regions and in rural/urban areas depending on the respective sizes and heterogeneity. It will not only ensure the quality but also reduce the over-all cost of collection of data. For achieving this, the minimum three-stage stratified sampling is necessary and in some cases it may have to be four-stage stratified sampling. However, the realistic problem with regard to re-allocation of investigators, the problem of local language for proper understanding and canvassing of schedules and the resultant additional cost involved in such operations may have to be taken care of and should be planned in advance carefully. Since it may not be feasible to shift large numbers of staff from one zone/region to another, the possibility of contract employment for data collection exercise in the respective regions may be assessed and due consideration for adequate training of such investigators may be taken into account. While it may necessitate some additional cost and administrative and financial problems, it may have to be managed for improving the quality and ensuring the reliability of data.

## 4. Efficiency of Sampling Technique

The estimation of efficiency of sampling procedure adopted is essential and it should invariably be included in the reports presenting the results of any round at the Central as well as at the State level for the various characteristics under study. This is not only necessary to establish the reliability and consistency of the results, but it can also help in comparing the results and quality of alternative techniques/ procedures and different agencies supplying related data. Besides, it can also give an idea about the reasonableness of sample size for future studies. The adequacy of sample size at the State level cannot be ensured in the absence of estimates of sampling error in the past rounds of sample surveys. Even for deciding the sample size and the sampling technique or the combination of techniques, the sampling error based on the pilot surveys under different techniques can and should be estimated. In the age of increasing competition from the private consultancy firms dealing in market surveys and other similar exercises, it is all the more essential to ensure the publication of estimates of sampling error along with the results. The NSSO has, therefore, to ensure estimation

and publication of the estimates of sampling error for each characteristic under various rounds. In case the sampling error is more than permissible for adequate reliability, the caution may be mentioned on the publication with respect to the reliability of the estimates in case the same is to be used. However, it may be desirable not to publish such results at the level the error is significant.

## 5. Importance of the National Sample Survey

There are various agencies conducting sample surveys of topical interest. These agencies include various industries, associations, federations and research organisations like the CSIR, the ICAR, etc. However, comprehensive information about the scope, coverage and adequacy of sample size and appropriateness of sampling methodology needs to be ensured. Although the information on the sample surveys conducted by various government organisations is being collected by the CSO, similar information from private organisations and autonomous scientific organisations like the CSIR, ICAR, FICCI, ASSOCHAM, etc., may also be collected and the views of the NSSO on the same, particularly with respect to sample size, sampling methodology, stratification, etc., may also be solicited. This is necessary to compare the reliability and quality of data published in the country. While the objective of sample surveys conducted by the NSSO is meeting the national requirements of policy decisions, for planning at the Central and the State level, the same for private and industrial associations may be different and not that exhaustive. As such, the results thrown up by the surveys conducted by these agencies may be considered with due care while utilising the same for any policy purposes. In view of this, it is necessary to increase the scope of the NSSO, so as to cover most of the surveys of topical interest besides the annual, quinquennial schedules of various subjects for the usual NSS surveys. The surveys required by smaller formations like cities, corporations, regions, divisions and sub-divisions, and for shorter durations like a quarter or a month, may be considered by the NSSO to supply quality information for different users besides the government. It may necessitate a significant increase in the resources, in terms of manpower as well as financial, for meeting the requirement of sample surveys and primary data at large level, so as to minimise the surveys by the interested parties, which may sacrifice some of the essential aspects and components of the concerned problem, required in larger public interest. It may also require enhancement in data processing facilities in terms of advanced modern computing equipment, other telecommunication facilities like internet, e-mail, etc. But it is necessary to bridge the gaps of the availability of various data for planning, particularly at the State, district and regional levels. It is also necessary to reduce the time-lag and improve the reliability and quality of information.

# METHODOLOGY OF NATIONAL SAMPLE SURVEY ORGANISATION AN APPRAISAL

A.K.Yogi[*]

## 1. Genesis of the NSSO[1]

On the recommendation of the National Income Committee headed by Prof. P.C.Mahalanobis with Dr. D.R.Gadgil and Dr. V.K.R.V.Rao as members, the National Sample Survey (NSS) was started in 1950 with a view to obtaining comprehensive and continuing information relating to social, economic, demographic, industrial and agricultural issues through large scale sample surveys. While instituting the NSS, the Government of India turned to the Indian Statistical Institute for providing technical direction to surveys. Thus, the work relating to finalisation of sampling design, schedules of enquiries, writing of instructions, training of field staff, processing of data and writing of reports was all entrusted to the ISI. The Directorate of NSS was created for the specific purpose of conducting field work. In July 1969 a three-man Committee comprising Shri B.Sivaraman, the then Cabinet Secretary, Prof. V.M.Dandekar and Prof. Raghuraj Bahadur, was set up by the Government for re-organisation of the NSS. The Committee recommended *inter alia* entrusting of all aspects of the survey to a single government organisation located in the Cabinet Secretariat and governed by a Governing Council.

The Government accepted the recommendations of the Committee and under the resolution of March 5, 1970 announced the decision to set up the National Sample Survey Organisation in the Deptt. of Statistics (which was then located in the Cabinet Secretariat but now a separate Deptt.) having four Divisions, viz. (i) Survey, Design & Research Division, (ii) Field Operations Division, (iii) Data Processing Division and (iv) Economic Analysis Division, later re-named as Data Analysis and Publication Division.

On the setting up of the NSSO, the Directorate of NSS was re-named as the Field Operations Division. The Divisions of Survey Design and Research and of Data Processing have been carved out of the employees deployed on NSS work in the ISI who were transferred to the government on account of transfer of NSS work from the ISI.

The expertise and experience of the ISI in general and Prof. Mahalanobis in particular provided the basic foundations for the NSSO. It is now almost five decades old. During these five decades, the NSSO has developed its own survey methodology suited to Indian situations. The experiences of the NSSO have also been amply made use of by other countries. In the year when the nation is celebrating its 50 years of independence it will be a tribute to the founders of the NSSO if the methodology adopted by the NSSO is reviewed and, if necessary, improvements are effected for meeting the future tasks of the 21st century.

The main objective of this paper is to present in brief an appraisal of the over-all methodology adopted by the NSSO with some suggestions for improvement.

## 2. Survey Organisation

The NSSO is a major instrument of the Government of India for collecting a variety of data on social and economic aspects of the population of the country. It carries out every year sample

---

[*] Joint Director (North Zone), FOD, NSSO, Jaipur. The views expressed in the paper are the author's own and not of the organisation to which he belongs.

surveys throughout the country in the fields of socio-economic, agricultural and industrial statistics. It has come to occupy a unique position in the country's statistical system as a major fact-finding body, in areas which are vital for development planning.

In the field of socio-economic statistics, the Organisation has taken over the National Sample Survey (NSS), the multi-subject integrated continuing household sample survey programme being conducted in the country since 1950. The NSSO covers different subjects of importance such as employment and unemployment, consumer expenditure, land-holdings, livestock enterprises; debt and investment, social consumption, demography, morbidity and disability, etc., through the household surveys. The socio-economic surveys are carried out in the form of rounds, the period of a round being normally a year. At present, the 54th round is on. The survey programme conforms to a cycle over a period of ten years, some topics such as land-holdings and demography being repeated once in ten years, and others such as consumer expenditure, employment and unemployment and unorganized segments of non-agricultural sector being repeated once in five years. Subjects of special importance are also accommodated in the remaining years or dovetailed with regular repeated surveys. This perspective permits the fulfilling of the felt needs for statistics for national planning and development as also the building up of time series on some of the important subjects at regular intervals to help measuring change over time.

In the field of agricultural statistics, the NSSO provides technical guidance to the States in respect of the Crop Estimation Surveys (CES), besides assessing the quality of primary work done by the State Agencies in area enumeration and yield estimation, under the scheme of Improvement of Crop Statistics (ICS).

In addition to these regular surveys, the NSSO undertakes *ad hoc* surveys like the Working Class Family Income and Expenditure Survey (1981-82), for the preparation of the weighting diagram for the Consumer Price Index Numbers for Industrial Workers and the Middle Class Family Living Survey (1982-83) for the Consumer price Index Numbers of Urban Non-manual Employees, etc. Regular price collection at rural and urban centres is also undertaken for use in the preparation of Consumer Price Index Numbers.

About the NSSO as a survey organisation, the United Nations Economic and Social Commission for Asia and Pacific has commented, "The first major attempt to integrate household surveys was made in India. The National Sample Survey of India was set up in 1950 which was designed as a multi-purpose Sample Survey system. It is a distinct body under the Deptt. of Statistics and is different from the Central Statistical Organisation. It has its own Governing Body under a non-official Chairman and has direct dealings with the State Bureaus of Statistics. It has a coordinating Office under a Chief Executive Officer and separate Divisions for Design and Survey Research, Field Operations, Data Processing, and Economic Analysis. Over the years, the NSSO has become one of the largest survey conducting organisations in the world. It has over 5000 persons working for it in about 170 offices spread in all parts of the country."[2]

## 3. Multi-Subject Enquiries

The Socio-Economic Household Surveys taken up by the NSSO can be termed as multi-subject/multi-purpose enquiries, whereas all other surveys are generally uni-subject enquiries. Multi-subject surveys represent an extreme form of integration. Theoretically, a multi-subject survey has been defined as a simultaneous investigation of several subjects, not necessarily closely related, in a single survey operation for the sake of economy and convenience. Further, data on different subjects need not necessarily be obtained from the same set of ultimate sampling units or even from the same type of penultimate sampling units. In NSSO surveys, these two aspects have been suitably adopted either independently or jointly for data collection. The multi-subject surveys of the NSSO, therefore, naturally have all the administrative, organisational and operational advantages associated with an integrated survey programme. In addition, they have the most important advantage arising

specially from the multi-subject coverage, namely, the opportunity to cross-tabulate and correlate information on different aspects of household life on the need for which there has been increasing emphasis.

There are, however, certain disadvantages inherent in such a methodology. A wide subject coverage necessarily involves (a) a very lengthy schedule or a set of schedules, (b) a multitude of concepts, definitions, code structures, procedures and even methods of enquiry such as physical observation, measurement, weighing, record-keeping and accounting, besides the usual interview methods, depending on the subjects covered, (c) long interviews, repeated visits, Investigator's fatigue and respondent resistance, (d) complexities of training and supervision, (e) ambitious tabulation plans aimed at making the best possible use of the wide variety of data collected, (f) strain on data-processing facilities, with inevitable bottle-necks, delays in processing and difficulties in the production of results, (g) compromises in sampling design and sample size, (h) compromises in concepts and definitions, (i) restrictions on the depth of inquiry on individual subjects and (j) on the whole, a sacrifice in quality of information.

Of late, quality concerns are on the increase and at the same time, field problems have multiplied due to better education, greater awareness and inadequate time available with respondents, etc. It would be desirable to have an in-depth assessment of the impact of the NSSO multi-purpose/subject survey programme with regard to quality of its results in the context of the present environment.

## 4. Sampling Design

The design of the Socio-Economic Surveys is a two-stage stratified one with villages in the rural sector and urban blocks in the urban sector as the first-stage units (f.s.u.'s) and the households as the second-stage units (s.s.u.'s). As per stratification, each State is divided into a few agro-economic regions which are generally groups of districts and, within each region, strata have been formed on the basis of, among others, the population of the districts. In the rural sector, the villages (f.s.u.'s) are generally selected with probability proportional to size (circular systematically) and in urban areas the f.s.u.'s, i.e. the urban blocks, are selected generally by circular systematic sampling with equal probability as the blocks are supposed to have been formed with approximately equal populations. The second-stage units are selected by SRSWOR or circular systematic sampling or any other convenient sampling technique, depending upon the survey. The States also participate in the programme generally with a matching sample. It is a matter of concern that the real purpose of the States' participation in the NSSO programme has not so far been realised. There is, therefore, an urgent need to look into this aspect.

A useful strategy that the NSSO has followed in the case of its household surveys with stratified two-stage design is to take a large sample of f.s.u.'s and a small sample of s.s.u.'s (households) per f.s.u. (because the contribution to the total variance from the between-f.s.u. variation is more important than that from between- s.s.u. variation). Further, to reduce substantially the variance within f.s.u.'s, methods such as stratification have been adopted for selecting s.s.u.'s. Although, such a strategy increases to some extent the time and cost in the field owing to greater work involved in travelling to the sample f.s.u.'s, camp setting and listing of households, it reduces the over-all sample size in terms of sample households and leads to reduction in enquiry time and demands on processing resources.

As stated earlier, in large-scale multi-subject surveys, compromises would have to be made in regard to the choice of the sampling design. It is, therefore, necessary at the design stage to ensure that the design adopted is a rational one for each of the subjects of enquiry. In a country like India where diversities exist, it is quite possible that the sampling design adopted may be very efficient for a particular subject and at the same time it may be inefficient for another. It would, therefore, be

appropriate to suggest that methodological studies in this field may be undertaken by the research institutions or by the NSSO itself.

## 5. Other Special Characteristics of NSSO Socio-Economic Surveys

**Sub-Sampling**[3] One of the most important contributions that has been made by the NSSO survey experience in the theory of sampling is the concept of sub-sampling, which enables quick estimation of sampling errors as also an assessment of the non-sampling errors. Since the inception of the NSS, the samples have been drawn in the form of independent sub-samples, so that each sub-sample provides a valid estimate of the population parameter. The present sampling design provides for two independent samples to be drawn for each of the Central and State samples, thus giving a total of four sub-samples. A comparison of the sub-sample estimates provides a quick and easy estimate of sampling error irrespective of the complications in the sample design. Further, the sub-sample estimates can be used to assess the quality of data and are helpful in studying differential agency bias, and in locating possible weak areas by a suitable design of the computation programme. They can also act as a check on processing errors.

In earlier years, analysis of sub-sample results and estimation of errors used to be undertaken, but of late, this practice appears to have been discontinued. Thus, in the absence of the estimates of sampling errors, it is difficult for the NSSO to face the following concerns:

(a) Statistically, on what basis, can reliability of NSSO survey results be ensured to the users?

(b) Statistically what is the basis to prove that results of Central and State samples differ and as such they cannot be pooled even now?

These concerns are raised just not for the sake of criticism, but answers to these are very important for the overall credibility of an organisation like the NSSO.

**Round and Sub-Rounds**[4] The concepts of round and sub-round are yet another contribution of the NSSO experience to the survey-conducting agencies. The concept of round was developed to delineate one survey from another over time. This facilitates planning and conduct of the surveys with some flexibility. As the economy and behavioural pattern of people in a country like ours depends to a large extent on agriculture, which is seasonal in nature, it was considered appropriate to make the round coincide with the agricultural year in order to take into account the seasonal fluctuations. Carrying out a continuing survey programme in the form of one year, also permits changes in the subject-mix and contents of subjects from round to round and enables re-orientation of the sample design to suit the subject-mix adopted in a particular round.

Further, instead of carrying out the work of data-collection of a round of a survey, over the entire survey period of one year in an amorphous manner, it was considered desirable to regulate the field work in a more structural manner by introducing the concept of sub-rounds. The entire survey period of a round is divided into four equal time periods, termed sub-rounds. The sample of primary stage units is drawn in such a way that each sub-sample provides a valid estimate of the parameter of interest. Efforts by the Field Operations Division were also made to ensure that the work of a sub-round is done within the specified period, permitting only marginal overlaps between sub-rounds. The NSSO experience of rounds and sub-rounds has proved and established the following merits of these concepts for the benefit of the survey conducting agencies: (a) ensuring even-spread of samples throughout the year; (b) sub-round estimates for the study of seasonal variation; (c) advance estimates for each sub-round as the survey proceeds, incidentally avoiding processing bottle-necks at the final stage; (d) progressive estimates for longer periods by pooling successive sub-round estimates. To take full advantage of these concepts, it is imperative that the survey starts on schedule and that the full time of one year is provided for the field work, if a survey round is of one-year duration.

**Reference Period** The choice of the reference period or the time period to which the data collected refer, is to be decided on the basis of the feasibility of obtaining reliable information, apart from other considerations like end-use. In general, a longer reference period tends to affect the quality of data through recall lapse, while a shorter period may increase the sampling error. The question of appropriate reference periods for different items has been the subject of research since the inception of the NSS. For consumer expenditure, the 4th and 5th rounds of the NSS experimented with two reference periods, viz. a "week" and a "month". Studies by Majumdar[5] and Mahalanobis and Sen[6] have indicated that a "month" is better than a "week" as a reference period from the point of view of both bias and sampling error. While this is being generally followed in recent surveys also, for some items like consumption of clothing and durable goods, a "year" is used as reference period in view of the pattern of purchases of such items over time. In morbidity surveys, where the reference period has a vital impact on the estimates of prevalence rates, it was found by Saha and De[7], after comparing reference periods of one, two, three and four weeks, that a fortnight would be the most appropriate reference period and a reference period of less than one week would lead to underestimation.

Another aspect that deserves mention is the practice of having a moving reference period which is one specified duration prior to the date of survey. For example, in a survey of one-year duration from July to June, the adoption of a one-month moving reference period would provide data for the period from June of the previous year to June of the current year with approximately equal representation of all the months in the obsevations made on the sample. This procedure, while giving an average picture for a year, also helps in the reduction of recall errors and fatigue.

The reference periods so adopted are operationally convenient for a vast country like ours. But supervision of field work shows that, in certain cases, the respondent (in the absence of records) is unable to provide data for the reference period in question but readily reports the information for a shorter period such as a week in case of consumption expenditure surveys. The Investigator is then compelled to inflate the figure suitably and record the entry for the prescribed reference period. The effect of such adjusted responses is yet another area of study. Further, with the availability of electronic data processing devices, a more flexible approach could be experimented. In the "instruction", a "favoured" reference period could be prescribed for an item but the schedule could provide columns for both "period" and data. The period would be the favoured reference period in all possible cases. In case, however, the respondent can give data only for a different period, the actual data and the corresponding period would be recorded. Instead of leaving the procedure for adjustment (such as y x 4 or y x 30/7) to the unknown judgment of the individual investigator, a standard procedure could be adopted during tabulation. The more important advantage of the change would be that it would be possible to study from the records the most common reporting period and, in the light of the study, to modify the favoured reference period in future surveys.

## 6. Sampling Frame

A pre-requisite for any sample survey is the existence of a suitable sampling frame with adequate information on relevant auxiliary characteristics. In earlier socio-economic surveys, sampling frame in the NSS consisted of census village and Enumeration Blocks (EBs) in rural and urban sectors, respectively. Census EB's were later found not suitable for the NSS and a special Urban Frame Survey (UFS) was conducted in 1960. Since then, the UFS blocks wherever available and census blocks where UFS was not done are being used as sampling frames. Since 1972, as a routine measure updating of old UFS frames is being done almost every five years. However, a few problems still continue in the frame for surveys such as (i) non-availability of same population census figures for a few States/regions, due to non-conduct of population census due to various factors and (ii) non-availability of uniform population census figures, leading to adoption of two sampling frames. This has its effect on the estimation procedure and on the survey results.

The problems referred to above are such that the NSSO can have little control on them. The NSSO experience in this regard, however, opens up yet another area of research and investigation.

For construction of CPI for rural agricultural labour, the NSSO, on behalf of the Labour Bureau at present collects prices from specified markets/shops. Since the base year frame has become almost 10 years old, the specifications of the items for which prices are being collected have undergone changes due to changes in consumption pattern. Also the shops/markets from where the prices are collected have not only changed but also gone out of existence. Similar is the case for middle class price collection data. Therefore, there is a need to up-date the frames for price collection schemes.

In the Annual Survey of Industries (ASI), the frame is prepared for a cycle of 3 years. This frame is built up from the information obtained from (a) the Chief Inspector of Factories as per Section 2m(i) and 2m(ii) of the Factories Act, 1948, (b) the Licensing Authority for *Bidi* and Cigar Workers, Condition of Employment Act 1966 and (c) the electricity undertakings registered with Central Electricity Authority. This frame is up-dated annually with a built-in system of obtaining every new registration/ deregistration of units.

The ASI frame, although revised every 3rd year, faces problems like (i) large number of entry of units every year through the annual supplementary frame owing to expansion of industrial economy, (ii) large number of changes in names, ownerships, location, employment size, products, etc., leading to wrong classification of units in the basic frame, (iii) inadequate and incomplete information with the concerned authorities, etc. All these factors keep the ASI frame fragile at a point of time. Such a fragile frame has its own effects on the results of the survey similar to that of an incomplete frame. No assessment of such factors has so far been made in the Central Statistical Organisation, which is the sponsoring agency for the ASI There is, therefore, an imperative need to look into the fragility of ASI frame to make it free from such inadequacies. Further, considering the need for district level estimates, a review of ASI methodology is considered necessary.

Insofar as the frame for agricultural surveys for Improvement of Crop Statistics (ICS) scheme is concerned, the list of villages under the Timely Reporting Scheme/Scheme for Establishment of an Agency for Reporting Agricultural Statistics and for General Crop Estimation Surveys (GCES), which is obtained from the State Agricultural Statistics Authority (SASA), forms the sampling frame. Generally, no frame problems as such are found for such surveys but of late, it has been observed in some States that the number of villages planned for sample check under ICS is so high that it accounts for all the villages under general crop estimation survey in some districts. On actual visit, it is also found that the villages planned under GCES do not grow the crop in reference. Even the crop in reference is not available in the additional villages. This results in loss of experiments for that crop. This suggests that while planning GCES, such problems should be looked into.

A large part of the non-agricultural activity, such as manufacturing, trade, transport, etc., is carried out as a household economic activity. Due to the unorganised nature of these activities, it is neither possible to get information on such activities through the normal administrative system, nor is a suitable frame for such units readily available. To overcome these problems, a pioneering scheme of Economic Censuses and follow-up surveys was taken up by the CSO and first Economic Census was conducted in 1977. Although initial problems faced in the frame based on the first economic census were rectified in the subsequent economic censuses, yet two chronic problems still persist and these are (i) difficulties in obtaining lists and EBs and their maps from the offices of the Registrar General of India, because such records are dumped in their godowns and stores after the census operations, (ii) improper demarcation and non-identifiability of EB's in the field.

Since the scheme of Economic Census (EC) and Surveys is the continuing scheme of the CSO, and its main objective is to provide adequate frame for the follow-up surveys, the CSO may, therefore, look into these problems and ensure for future that lists of EB's are procured by them soon after the Economic Census is over from the Census-conducting agencies and make them available to the FOD for follow-up surveys.

## 7. Schedules and Instructions for Field Staff

The schedules of survey are prepared with due care for structure, sequencing of items and wording. They also admit of detailed on-the-spot probing unlike in a questionnaire approach. Usually in the schedules of the NSSO, general particulars relating to households are covered first, followed by the core items of the subject of enquiry, in a logical order. Sensitive items are covered at the end after establishing a good rapport with the respondents. Provisions also exist in the schedule for summarising the observations to enable quick tabulation, etc. For the conduct of each survey, a comprehensive set of instructions for field workers is documented indicating in detail how to collect and record the data in a pre-formatted schedule. Schedules and respective sets of instructions, besides other necessary materials like standard classification of Industry and Occupations, the glossary of technical terms, etc., constitute the basic survey material for the primary field worker.

No doubt, the preparation of such survey material for the field personnel greatly helps in controlling non-sampling errors. Yet, it has been found that the size and structure of schedules of NSS enquiries are quite heavy and bulky and suffer from the drawbacks enumerated earlier in Section 3 of this paper. There is, therefore, a need to make a serious attempt for shortening the schedules of NSS enquiries to make them truly suitable to the prevalent field conditions.

## 8. Modernisation and Use of Palmtop Computers

In the current Five-Year Plan, the Deptt. of Statistics has initiated some programmes for modernisation of the statistical system in the country. One of these programmes is the introduction of Palm Top computers for data collection in FOD. No doubt, the Palm Top computer is a useful and an efficient device. Yet, the experiment so far in FOD reveals genuine problems for its effective usability in the field under the prevailing situations. The last two years' experience of palm-top Computer of FOD suggests that for taking full advantage of this device for achieving efficiency and economy in resources like time, cost and manpower, the following aspects merit consideration:

(i) The survey schedules need considerable reduction in size;

(ii) dependency on the manufacturer of the device has to be overcome for development of software;

(iii) software packages be centrally developed and then made available to the field offices;

(iv) linkages and co-ordination aspects between FOD and DPD/Computer Centre need strengthening; and

(v) FOD may not be assigned processing (even partially) jobs unless the present field set-up is re-organised to cater for such activities.

## References

1. NSSO (FOD), North Zone, Jaipur: Unpublished Training Literature.
2. UN Economic and Social Commission for Asia and the Pacific, *Report on Household Surveys in Asia - Organisation and Methods* (1981).
3. Murthy, M.N. *Sampling Theory & Methods*, Statistical Publishing Society, Calcutta, 1967.
4. *Ibid.*
5. Majumdar, D.D. *Reference Period in Family Budget Enquiries*, NSS Technical Paper No. 43/70/1.
6. Mahalanobis, P.C. and Sen, S.B. "On some aspects of National Sample Survey", *Bulletin of the International Statistical Institute, 34,* Part 2 (1954).
7. Saha, R.P. and De, A.K. *Some Aspects of Morbidity Surveys in the Indian National Sample Survey,* NSS Technical Paper No. 28/68/14 (1968).
8. UN *ESCAP op. cit.*

# RESTORING THE BALANCE

## A.M.Goon[*]

While talking of the Indian Statistical System, I shall take a wide view and include in the system the application of statistics in the various branches of science, especially in the work of socio-economic development of the country, and also the work of teaching and theoretical research in statistics that is done in the country's universities and institutes, besides the system of collection and compilation of data through the diverse governmental agencies. During what may be called the *Mahalanobis era* of Indian statistics, all three wings of the system grew and prospered. In more recent times, however, applied research and the system of official statistics are not getting as much attention as they deserve. A certain imbalance has thus developed in the system, which will ultimately be detrimental to the statistical profession. I propose to consider the history of development of statistics in India and also the nature of the imbalance. Some remedial measures will be suggested towards the end of the discussion.

### A Glorious Era

Think of the time when Professor P. C. Mahalanobis started working on statistics from a small room in Presidency College, Calcutta, where he was employed formally as a teacher of physics. That was virtually also the time when statistics in India started its journey. Initially, the Professor handled some problems of a practical nature: analysis of some Calcutta University examination results, determining the relative efficiency of different types of vaccine, studying the effect of weather conditions on the incidence of diseases, analysing anthropometric data, etc. But soon enough he started devoting attention to theoretical research and gathered around him a group of young researchers like R.C.Bose, S.N.Roy, S.S.Bose and D.B.Lahiri. When he built up the Indian Statistical Institute in 1931, he had primarily in view theoretical research in statistics. The Institute gradually grew in size and also in the range of its activities. The ISI scientists' contributions in the fields of experimental designs, survey sampling, inference and multivariate analysis received worldwide recognition. When the ISI was shifted from Presidency College to its sprawling campus at Baranagar in the northern outskirts of the city, the Professor was already conscious of the immense possibilities of the new discipline as a harbinger of new methods of research and new ideas in all branches of knowledge. Hence one finds in the ISI a separate unit for almost every branch of science that might call for application of the statistical methods of research. Indeed, he provided the ISI even with a separate unit for linguistics! At the same time, he was alive to the fact that the new discipline could not strike root in the Indian soil unless it concerned itself with the task of national reconstruction, unless it addressed itself to the immediate socio-economic problems of the country. Hence his interest in large-scale sample surveys, controlling floods in the rivers of eastern India, determination of national income and, above all, in economic planning. But he soon realized that efforts towards the socio-economic development of the country could not bear fruit unless the database of the Indian economy was strengthened.

His proximity to and warm personal relationship with the leaders of the Indian independence movement, especially Jawaharlal Nehru, enabled the Professor to bring about the changes that he deemed necessary. He had met Nehru a number of times before 1940 when Nehru had come to Santiniketan to visit poet Tagore. But it was in 1940 that he had the opportunity of spending a day with Nehru at the latter's residence (Anand Bhavan) at Allahabad and having discussions with him on the importance of planning and the major *rôle* of statistics in planning.

[*] Retired Professor of Statistics, Presidency College, Calcutta

The President of the Indian National Congress (Subhas Chandra Bose) had earlier (in 1938) set up the National Planning Committee with Nehru as chairman. Nehru raised the question of having adequate and reliable statistics at the 3rd session of the Committee held in May 1940. Later, in March 1946, Nehru came to Calcutta for a week's stay to preside over the session of the Indian Science Congress. He stayed for a few days at the Baranagar residence of Professor Mahalanobis, who was at that time General Secretary of the Indian Science Congress. Professor Mahalanobis took him round the ISI and apprised him of the wide range and importance of the work being done there. The importance and indispensability of statistics in planning any future programme of development of the country was explained to Nehru. Professor Mahalanobis even suggested that some bright young members of the National Congress be sent to the ISI for training in statistics. The very next month, Nehru sent Pitamber Pant to the ISI in response to the suggestion. (Pant's *rôle* in the work relating to the first few Five Year Plans is now part of the history of planning and statistics in India.)

Quite in the fitness of things, shortly after the attainment of the country's independence, Professor Mahalanobis was made Honorary Statistical Adviser to the Union Cabinet. Thanks to the keen interest of Nehru and C.D.Deshmukh, the then Financial Adviser to the Government of India (who became Minister of Finance in 1950), Mahalanobis could set up a Central Statistical Unit (CSU) in the Cabinet Secretariat and a small office of the ISI in New Delhi. Pitamber Pant and the Institute staff helped in running the CSU for about two years until the CSO was set up in 1951 for coordinating the activities of the statistical offices of the Central government and also those of the State Statistical Bureaus, which were set up in all the States. In 1950, the NSS was started for collecting data on social, economic and demographic features of the country on a continual basis.

In 1960, the ISI was recognized by an Act of the Parliament as "an institution of national importance", paving the way to its development as a centre of excellence with strong financial backing from the Central government. The Act empowered the ISI to start courses of study leading to the degrees of Bachelor of Statistics (B.Stat) and Master of Statistics (M.Stat) and also to confer the degrees of Ph.D. and D.Sc. in statistics.

The work of Indian statisticians soon came to earn worldwide acclaim. Many of them were offered faculty positions in prestigious universities in the USA, UK and other countries. Many others came to occupy key positions in various United Nations agencies. Many developing countries sought the assistance of the Indian government in building up their own statistical systems.

It was indeed a glorious era for Indian statistics.

**Remarkable Balance**

The balance that I referred to right at the start of this discussion was largely due to the outlook on statistics of Professor Mahalanobis. While considering any new statistical technique or any new model of analysis, he would always keep in view possible applications instead of being attracted by the underlying mathematics. He wrote: "Statistics, like engineering, requires all the help it can receive from mathematics, but... (statistics) can never become a branch of mathematics." He was more emphatic when he said at a later stage that "mathematical statistics as a separate discipline cannot simply exist". To him statistics as a discipline was useful only for its possible applications. He said : "Statistics is essentially an applied science, its only justification lies in the help it can give in solving problems." Gradually, he came to believe, with R.A.Fisher, that "Statistics is a key technology meant to increase the efficiency of human effort in the widest sense."

When the ISI was still housed in a part of Presidency College, Calcutta, for some years the departments of statistics of both Calcutta University and Presidency College were run by the members of the staff of the ISI working on a part-time basis. The departments used to hold their classes in the same part of Presidency College where the ISI was situated. Occasionally, Professor Mahalanobis would hold unscheduled joint classes for the (Honours) students of Presidency

College, (M.Sc.) students of Calcutta University as well as ISI trainees (mostly serving in different government departments). These students would sit together and listen to Professor Mahalanobis speaking on some of the basic things of statistics: questionnaires and schedules, various modes of data-collection, checking data for possible errors, taking repeated measurements to develop the ideas of location and dispersion, etc. I mention all this to emphasise that Professor Mahalanobis wanted his students to remember that the collection, analysis and interpretation of data were the ends of any statistical study. The tools of statistics were important only insofar as they could achieve these ends and not for their mathematical elegance. Regarding the teaching of statistics, his views were quite clear. In his opinion, "the teaching of statistics should be looked upon as something analogous to the teaching or training in engineering or medical sciences". During a visit to a US university shortly before the B.Stat. and M.Stat. courses were started at the ISI, he said in course of a lecture, "in India, in the Indian Statistical Institute, we hope to be able to introduce very soon, full professional courses leading to the degrees of Bachelor of Statistics and Master of Statistics, the degree of Bachelor of Statistics being exactly comparable to a professional degree in engineering or medicine".

I firmly believe that the statistical profession flourished in India and received encomiums from all around during the *Mahalanobis era* because of this outlook and because of the healthy balance among the three wings of the statistical system.

## The Present Scenario

The golden era of Indian statistics virtually ended with the death of Professor P.C. Mahalanobis. Things remained undisturbed till Professor C.R. Rao remained at the helm of affairs. But his increasing pre-occupation with mathematical statistics and eventual departure from India disturbed the balance that I have spoken about in the earlier paragraphs. Indian statistics is now without a proper direction and is very much like an unguided ship drifting on the high seas.

High-quality research in the field of mathematical statistics continues to be done by Indian statisticians. Their research papers appear regularly in prestigious international journals. Some are still invited to occupy faculty positions in universities in the USA, Canada, UK and Australia. Many more frequently go abroad to attend seminars and symposia, where they are heard with respect.

But surely there is now much less significant work in the field of applied statistics. Papers do come out in the fields of experimental designs and survey sampling. But they are, by and large, concerned with construction of designs or formulae for estimates, and for estimates of their standard errors, for some variants of the already available types of sampling. Nobody takes the initiative to embark on large-scale sample surveys of the type undertaken by Professor Mahalanobis for estimating the area and yield of major crops in Bengal and Bihar or a survey like the UP Anthropometric Survey undertaken by him in collaboration with Professor D.N. Mazumdar of Lucknow University. Statisticians are seldom consulted these days in matters of socio-economic planning. There has been no statistician as a member of the Planning Commission for a long time now.

But I suppose the worst casualty has been the system of official statistics. Statisticians working with various ministries or departments seem to have lost all initiative for undertaking any innovative projects. Even their routine functions (like data collection on the same items year after year and getting them published in official handbooks or abstracts) are done in a lackadaisical fashion. The authorities, too, seem to attach little importance to the work done by the government statisticians.

Some time ago, I had an occasion to look up some old volumes of the *Annual Statistical Abstract* brought out by the Bureau of Economics and Statistics of a major State of the country. The

publication came out very regularly since 1944 till 1974. But then there was a single issue for 1975 and 1976 and again a single issue for the entire period from 1978 to 1989. This much for the timeliness of the published data! On the question of reliability of the data, I may give here from the 1975-76 issue part of a table which is supposed to present district-wise data on sex-ratio at birth for the years 1971 to 1974[*] :

| District | 1971 | | 1972 | | 1973 | | 1974 | |
|---|---|---|---|---|---|---|---|---|
| | Male | Female | Male | Female | Male | Female | Male | Female |
| (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) |
| All distts. | 52.50 | 47.50 | 6.8 | 6.2 | 6.9 | 6.2 | 6.6 | 6.0 |
| Distt. A | 52.87 | 47.13 | 3.7 | 3.9 | 4.3 | 4.0 | 3.4 | 3.2 |
| Distt. B | 53.04 | 46.96 | 9.2 | 7.9 | 9.1 | 8.0 | 9.0 | 8.5 |

Columns (2) and (3) surely give the percentage of male births and the percentage of female births (for 1971); but what do the figures in the subsequent columns stand for? Readers may feel amazed that such a table could be cleared by the Bureau of Economics and Statistics for publication. What is more, to my knowledge, the errors have gone undetected to this day and have, in fact, reappeared in a later issue of the publication.

Such serious lapses on the part of the Bureau only point to an utter apathy in the minds of the qualified statisticians employed in the Bureau. In fact, many of our statisticians serving the official statistical system feel that their work is not valued by the theorists in the universities and institutes. Thanks to the way statistics courses are now drawn up in the universities and institutes, they seem to develop an aversion to data and an aversion to figure work of any kind whatsoever.

In the present situation, therefore, applied research is at a discount. Even if some researchers take interest in research of this type, they may feel discouraged by a lack of sufficiently up-to-date and sufficiently reliable data in the official statistics sources. Virtually the only type of research that gets the support of our leading academics is theoretical research - research with a sufficiently mathematical content.

## Remedial Measures

The Indian Statistical System cannot be left in this state of drift. For one thing, the need for timely and reliable data cannot be over-stressed for the smooth functioning of the government. There is also a paramount need for the application of statistical methods for research in the sciences, including the social sciences. If this is neglected, even theoretical research will lose relevance to the socio-economic needs of the country. What is more, if the needs of the country are ignored, if statisticians can offer no help to the statesmen or researchers in the other sciences, the very future of the statistical profession may be in jeopardy.

It is in the interest of the statistical profession that our academics should come forward to restore the balance of the statistical system. By this I mean that while people with a mathematical bent of mind and the necessary talents should continue making contributions to mathematical statistics, there should be more of application-oriented research in statistics. Besides, people running the official statistical system should be encouraged to improve the system so that timeliness, reliability and relevance of the collected data may be ensured.

A change in the attitude of the leading lights of the profession is clearly called for. They should begin by making changes in the statistical curricula in the universities and institutes. In this

---

[*] Source and identity of districts are not disclosed to avoid possible embarrassment.

connection, the following passage from an unpublished manuscript of the late Professor S.N. Roy deserves to be quoted. To bring into focus the true *rôle* of statistics in the scheme of things, he wrote : " ....... we have first of all a sector of the physical or phenomenal (or even conceptual) universe that presents to us certain questions posed with an eye to knowledge or to action ..... Next, we have a plan of experimentation or taking observations and a logico-mathematical model, both interlinked, the two together being regarded hopefully to be ........ an appropriate formal tool or means for answering the original questions. Third, from the data now formal answers in the sense of inference are sought within the logico-mathematical framework; and fourth, the formal answers are now sought to be related to the original questions or purposes." What Professor S.N.Roy meant was that people (administrators, business managers or researchers in the various sciences) would come to a statistician to seek his assistance in solving their problems. Generally, the questions asked by these people would be framed in a vague, imprecise manner. The statistician would have to reformulate each question so as to place it within a probabilistic framework; i.e., he would have to choose a p.. ·ability model from among a number of possible alternatives. He would then collect the necessary data from the relevant field either through experimentation or through survey sampling or, playing a more passive *rôle*, take the data as they appear in a chronological order (as happens in the case of time-series data). He would at the next stage analyse the data in accordance with his chosen probability model. Finally, he would have to interpret the findings to the people who initially asked the questions, and this he has to do in a language intelligible to those people.

If these be the functions that a statistician is expected to perform, then the place of mathematics in a statistical curriculum can be seen in its proper perspective. A knowledge of mathematics is essential for grasping the implications of the probability models and the theoretical bases of the tools used in tackling problems. But if the mathematics becomes highly abstract and more time is devoted in a statistician's training programme to mathematics than to the collection and scrutiny of data, the rationale behind the statistical tools, their utility as well as limitations, and to the correct interpretation to be put on the findings of data analysis, then the training imparted is bound to be lop-sided. This is exactly what is happening in our prestigious seats of learning. That the results have been all but disastrous is not surprising. But I, for one, feel amazed at the utter lack of understanding of such simple tools as, say, the correlation coefficient, the paired $t$-test and Fisher $t$-test and the various applications of the Pearsonian $\chi^2$ statistic shown by M.Sc.s and Ph.D.s in statistics when they appear before interview boards. The fact is that during their training, they have learnt to attach more importance to the needed mathematical equipment so that they can reel off the various theorems, properly stating the underlying assumptions. But relating the theorems to practical problems through the use of commonsense is something to which they are not accustomed. Hence the aversion to data and to data analysis that we find among the young graduates.

In the interest of the statistical profession, the imbalance in the Indian Statistical System has to be removed, and the work of reform has to start with a proper re-orientation of the courses offered by our universities and institutes. For one thing, the students must at the first stage be told that data-collection and data-analysis are what they will be expected to do and the various items of their training programme are meant to equip them with the necessary expertise. Hence, for instance, their training programme must start with a course in descriptive statistics and not with one on mathematical analysis or probability theory; besides, what is called "practical work" must form an integral part of the course.

I have made the above suggestions since I feel alarmed at the present state of the Indian Statistical System and yet believe that we can put it back on the rails if we proceed with imagination. What is needed is leadership with the necessary will and dynamism and with a comprehensive approach to statistics. Professor Mahalanobis could achieve what he did because he had vision and dynamism with a holistic approach to statistics. Indeed, nurtured in the ambience of the Bengal Renaissance, he had a holistic approach to life itself.

# QUALITY LEARNING IN STATISTICS

Gopal K. Kanji[*]

## 1. Introduction

Statistical learning can be defined as a process in which individuals can change their attitude to adopt a continuous development of basic knowledge and skills in pursuit of total professionalism.

The essential feature of an effective learning process is constant updating and continual feedback. Thus, the effective learning process in an organization should be focused around the quality of feedback provided by the organization. Effective action must be organised around a range of system and procedures to accomplish the goal. Learning to learn involves the continuous development of various strategies and skills that support the process of learning in many different contexts of statistics.

The basic requirement of any effective learning of statistical process is, therefore, the desire to learn the skills, to implement them and to practise them in an appropriate context. However, continuous learning requires a sustained interest in learning over time and relates to the improvement in learning ability, which is independent of the content being learned.

Experience shows that developing ideas for quality improvements is an investment which gives the highest return on resources. We believe that educational institutions have many opportunities for improvements, which should be fully utilised. If employees (including teachers) and students are educated and trained in total quality in statistics they have the joint capability to utilise these lost opportunities. Education and training is only one, albeit necessary, condition for the involvement of an educational institution's employees and students in continuous improvements. However, continuous improvements also require leadership, which is the foundation of total quality in statistical education.

Talking about the product of education, we agree with the definition by Bower (1994); "The product of education is a person empowered to educate herself or himself - educating person not an educated person." In this modern world, a person can never be truly educated. What a person learns within a limited time period, say 4 years, will inevitably be more or less outdated within a few years. Put another way (Bower):

The product of education is not a product at all; it is a *process* of never ending education. To be empowered in this way requires that the student gain from education:

(1) Knowledge of how to continue learning and developing for the rest of her/his life.

(2) Skills in continuing to learn and develop.

(3) A state of mind and being that enables lifelong learning, that sees and feels self-guided learning to be natural, doable and fun.

(4) A strong drive, a will, to continue learning and developing her/himself.

---

[*] Professor, Centre for Quality and Innovation, Sheffield Hallam University, City Campus, Computing and Management Sciences, Sheffield, UK

The key word in relation to continuous improvements is learning. In order to communicate this to his audience/readers, Deming changed the name of the improvement cycle (the Deming cycle) from plan-do-check-act (PDCA) to plan-do-study/learn-act. In the check phase of the improvement cycle, you have to study the results in order to understand what were the causes behind it. This learning process is the most important part of the continuous improvement process. Therefore, we will discuss the learning process.

The next sections discuss the goal, strategy, various learning processes, **advantages and** disadvantages of the continuous development of a quality professional.

## 2. Continuous Learning

There is a common misconception that learning in relation to working life at the professional level ceases when full-time education comes to an end. This may be at the first degree, master's or doctoral level. The reality could not be further removed from this.

There are a variety of learning approaches, but the major ones include:

- studying for examinations for professional membership, e.g. Applied Statistics;

- part-time study for master's and doctoral degrees by study and research, e.g. M.Sc., M.Phil. and Ph.D. in Applied Statistics;

- learning at work.

It is this last approach which is the focus of attention in this paper. Training in the workplace can follow many patterns: it can be *ad hoc* and uncoordinated, with much of the responsibility put on the individual, or it can be planned with some assessment of efficacy. This paper looks at one learning model.

## 3. Common Learning in Statistics

It is commonly assumed that statistical learning is strongly associated with the teaching process, when information is given by a teacher. The application of the assimilated information is then applied to appropriate tasks. This is not now thought to be the case. Most real learning is achieved via experience. This is 'on-the-job' experience. There is a common learning cycle (see Fig.1).
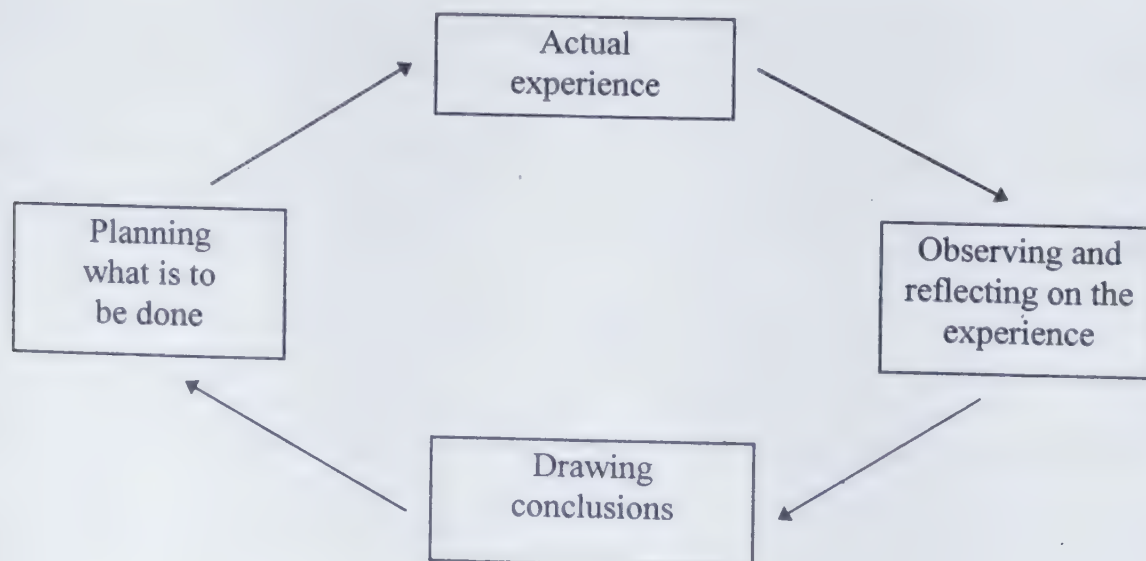


**Figure 1**. *Common learning cycle.*

The essential feature is that of feedback; action is taken, the effect is noted and appropriate modification of behaviour follows. If real learning is to be achieved, then all stages of this cycle must be experienced.

The development of this statistical learning cycle as a model of good practice for effective learning is the focus of this paper. There are many issues relating to this learning cycle which can be established in an organisation, but we must first look at how to start the learning cycle process.

## 4. Effective Learning in Statistics

There is an obvious need to ensure effective learning in statistics. If learning is not treated seriously, then there will be adverse repercussions for the organisations concerned. An effective learning strategy is likely to encompass the following:

- understand how any why people learn;

- promote a healthy learning environment;

- identify the individual's learning needs;

- prepare a learning plan with agreed objectives;

- promote learning opportunities;

- evaluate learning outcomes.

## 5. Quality Learning

We have talked a little about how adults learn through the 'learning cycle'. If people are really to learn, then each stage of the cycle must be travelled. There are many other factors involved in this process (some of which we will see later). What we need to do is to place this statistical learning cycle within the organisation.

Within an organisation, there may be many equivalent cycles for areas of development. Here, each cycle needs to be complete and unbroken for effective development to be possible. Take, for instance, the continuous statistical learning process, which includes all organisational activity. The continuous statistical learning cycle or continuous improvement cycle (Deming), which helps to improve the statistical activity of the organisation, can be depicted in Fig.2.

In general, statistical learning is a continuous process that can be stopped anywhere in the learning process of the customer-and-supplier chain system.
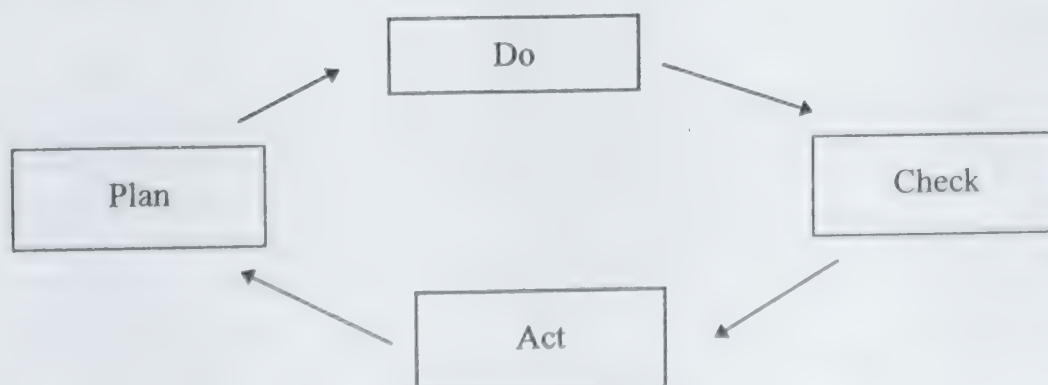


**Figure 2.** *Continuous improvement cycle (PDCA) (Deming)*

In the Deming cycle, 'plan' defines the learning process, which ensures documentation and sets measurable objectives against it. The 'do' executes the process and collects the information and statistical knowledge required. The 'check' analyzes the information in a suitable format. The 'act' obtains corrective action using statistical learning techniques and methods, and assesses future plans. At the end of each cycle the process is either standardized or learning targets are adjusted based on the analysis, and the cycle continues.

The continuous improvement cycle and the learning cycle can be superimposed, i.e. a planned, statistical-improving strategy. We now need to consider the learner and how his/her needs should be considered and also look at the preparation of a learning plan.

## 6. Learning Plan

Before a plan can realistically be drawn up, both the learning needs of an individual and the desired learning outcomes for an individual need to be considered.

To identify the learning needs of an individual, the knowledge, skills and attitude required for a particular task or function to be performed successfully must be established. Against this list should be considered the knowledge, skills and attitude already possessed by the individual. This second list will probably be achieved by consideration of career history and known work record. The difference between these two lists is the learning need (or learning gap). Part of the process of the identification of learning needs should be open discussion with, and agreement of the individual concerned.

Learning outcomes fine what (which tasks) an individual will be able to perform and how well he/she will be capable of performing these tasks as a result of the learning process. There should be a list of required outcomes in terms of behaviour and performance. The standard required should also be established, so that the effectiveness and the amount of learning can be measured. There should be a time horizon for this learning process.

The learning plan is thus, clearly, a bringing together of learning needs and learning outcomes. This plan, which is best formalized in writing, should be jointly agreed to by both the individual and the trainer/supervisor. The plan should follow the learning cycle and so will consist of a list of desired outcomes together with a list of activities necessary for the desired outcomes to be achieved. There may well be resource and specific training needs identified at this stage. This can be viewed as the planning stage of the cycle. Once the learning experience has been completed, some reflection/assessment should follow. It is hoped that there will have been appropriate consideration of the measurement of the desired outcomes. The final essential stage of the learning cycle is that conclusions based on the reflection and/or assessment are fed into reformulating and refining the next planning stage of the learning cycle. The process goes on and on: continuing and continuous learning.

## 7. Conclusions

In this competitive business world, the traditional approach to quality is no longer desirable. The basic problem that prevents an industry from becoming a quality statistical learning organization is connected with management and process. The author believes that the best way of overcoming the organizational quality problem is through the practice of TQ. To practise TQ, it is necessary to identify the basic activities affecting the organizational management and process. The approach to deal with these activities affecting management and process is, therefore, through education, training, research and quality learning in statistical activity.

Further, the continuous development of a statistical professional can be placed within the framework of a continuous improvement learning cycle. The implementation and monitoring of this

learning model can be evaluated with the cooperation of a panel of employing organizations in the context of the professional experience.

---

## Reference

Bower, M J (1994). "The paradigm shifts required to apply TQM and teams in higher education" in: H. Costin (ed.): *Readings in Total Quality Management*, The Dryden Press, New York.

# EXPORT SWING : A STUDY OF CHANGE IN INDIA'S MARKET SHARE IN THE ECONOMIC BLOCS

S. K. Nath[*]

*Abstract: It is well-known that the share of developing countries in the world market is very little and so is the case of India. But with the emphasis given on the openness of the economy in the International market, it is expected that India's performance of market share in world trade would improve. This paper studies India's market share according to various economic blocs of the world and makes use of an innovative approach to measure their swing.*

## Introduction

This paper aims at a study of India's changing scenario of market share in the world merchandise trade according to economic blocs of the world. For this, Hickman's (1979) formulation for measuring "changes of market share" and "changes in import share" was adopted. For the developing countries, a study of changes in market share is quite important. Such a study will be more useful if one could know the regional distribution of such changes.

For further interpretation of the results obtained on "changes in market share" and "changes in import share", we have measured their swing against defined desired level of performance.

## Data Sets

The database for this study will consist of the following data sets covering the period 1985 to 1994:

| Srl. No. | TYPES OF DATA SET |
|---|---|
| 1. | Total merchandise exports /imports of India |
| 2. | Countrywise exports of India |
| 3. | Imports of all the countries of the world |
| 4. | Economic regionwise exports of India |
| 5. | Imports of the various economic regions of the world |

## Data Source

In this study, we have used calendar-wise data since the partner countries' data are available according to calendar years only. In order to generate matching data of India, India's data were converted from the financial year of April-March to the period January -December using suitable methodology. We had converted India's external trade data into US $ using the Indian official exchange rates as available from the Reserve Bank of India from its monthly journal the *Reserve Bank of India Bulletin*. Data of the partner countries have been compiled from the various issues of the *International Trade Statistics* brought out by the United Nations. The domestic data on external trade have been compiled from the *Monthly Statistics of India* (March issues) as brought out by the DGCI&S, Ministry of Commerce, Government of India.

---

[*]Joint Director, Industrial Statistics Wing, CSO, Calcutta. The views expressed by the author are his own and not necessarily those of the organisation to which he belongs.

## Economic Blocs

For this study the economic blocs of the world as defined by the Government of India have been used. At present, there are 193 countries with which India has regular trade and there are sixteen(16) such economic blocs. The data of the 193 countries have been grouped in order to generate the economic blocwise data sets.

## Methodology

For the algebraic representations of Hickman's model and for modelling for export swing, we introduce the following definitions:

$X(i,j,t)$ = i-th country's exports to the j-th economic bloc during time $T = t$.

$M(j,t)$ = total imports of the j-th economic block at time $T = t$

$$= \sum_i X(i,j,t).$$

$T[X(t)]$ = total exports of India during time $T = t$

$W(t)$ = world exports during time $T = t$.

Let us further define the following:

$\alpha(j,t)$ = Market share of Indian goods in the j-th economic bloc

$$= \frac{X(i,j,t)}{M(j,t)};$$

$\beta(j,t)$ = import share of the j-th economic bloc with respect to the world share

$$= \frac{W(t)}{M(j,t)}$$

Thus, the total export is

$$T[X(t)] = W(t)[\ \sum \{(\alpha(j,0).\ \beta(j,0)) + (\alpha(j,0).\ \delta\beta(j,t)) + (\delta\alpha(j,t).\ \beta(j,0)) + (\delta\alpha(j,t).\ \delta\ \beta(j,t))\}],$$

where $\delta$ is an operator to generate the time difference.

While the first term of the above expression is a constant, the second term measures the changes in the import share, the third term measures the changes in the market share and the fourth term is a constant. Thus our interest lies in the derivation of the third term which we define as MS(j).

## Behavioural Pattern of MS(j)

In this study, we shall measure the fluctuations of MS(j) through measurement of its upswing and downswing. For this purpose, let us imagine two STATES, namely, S(u t) and S(d,t)

where

$S(u,t)$ = upswing during time $T = t$   and

$S(d,t)$ = downswing during time $T = t$,

such that

$MS(j,t) < S(d,t)$  implies   $MS(j,t) < S(u,t)$  and

$MS(j,t) > S(u,t)$ implies    $MS(j,t) > S(d,t)$.

We are interested in studying the situations when

$MS(j,t) > S(u,t)$  when   $MS(j,t-1) < S(u,t-1)$  is TRUE :        SITUATION-I

$MS(j,t) < S(d,t)$  when   $MS(j,t-1) > S(d,t-1)$  is TRUE :        SITUATION-II

Obviously, the two situations are mutually exclusive by construction. For a particular value of "t" we define a positive swing of market share of time lag 1 if Situation-I takes place. Similarly, we define a negative swing of market share of time lag 1 if Situation-II holds good. We can also define positive swing and negative swing of time lag 2 or above depending on the number of years when such swing takes place.

### Estimation of S(u,t) and S(d,t)

Estimation of the two STATES will depend on what importance we intend to attach to measurement of swing. Since in many cases it will not be possible to measure the actual probability distribution, we measure the STATES using the relations below:

$S(d,t)  =  E(MS) - k.\sigma$

$S(u,t)  =  E(MS) + I.\sigma$

where $E(MS)$ stands for the expectation of $MS(j,t)$ over the predefined successive values of $T = t$ and $\sigma$ stands for its standard deviation, k and I are two constants. Without the knowledge of the distribution of market share, one can derive the value of probability under two different situations using the Chebycheff lemma taking $k = 1$. Since our interest is to see that market share should not go below the average market shares, we shall put the value of the constant $k = 0$. For computation of $S(u,t)$, we tried for stronger to stringent and ultimately fixed it at $I = 2$ and computed the values of $S(u,t)$ accordingly.

In order to study the swing pattern, we have calculated the swing frequency and also the swing ratio. By swing ratio we mean the ratio of swing frequency to the number of outliers.

A similar analysis has been taken up for studying "Change in Import Share".

**Findings**

Table I    Study of upswing and downswing of change in India's market share
in the world merchandise trade

| Economic blocs | Upswing behaviour | | Downswing behaviour | |
|---|---|---|---|---|
| | Swing frequency | Swing/ outlier | Swing frequency | Swing/ outlier |
| East African Community | 1 | 0.33 | 0 | 0 |
| Organisation Commune Africanine El Malagache | 1 | 0.2 | 0 | 0 |
| Organisation of African Unity. | 1 | 0.17 | 0 | 0 |
| Other African Countries (OAC) | 1 | 0.25 | 0 | 0 |
| North America | 1 | 0.50 | 0 | 0 |
| South America | 1 | 0.20 | 0 | 0 |
| Caribbean Common Market | 1 | 0.50 | 0 | 0 |
| Central American Common Market | 2 | 1 | 0 | 0 |
| Other Central American & Caribbean Countries | 2 | 0.25 | 1 | 1 |
| Economic & Social Commission for Asia & Pacific | 1 | 0.33 | 0 | 0 |
| Economic Commission for West Asia | 2 | 0.40 | 0 | 0 |
| Other Asian & Oceanian, Antarctic and Arctic Region Countries | 2 | 0.50 | 0 | 0 |
| European Common Market | 2 | 0.50 | 0 | 0 |
| European Free Trade Area | 1 | 0.33 | 0 | 0 |
| East European Countries | 1 | 0.5 | 1 | 0.5 |
| Other European Countries | 1 | 1 | 0 | 0 |
| All Regions | 2 | 0.5 | 0 | 0 |

From Table I above, we find that the behaviour of India's "change in market share"(CMS) abroad has a somewhat upward bias except for the region Other Central American and Caribbean countries (OCACC). No downward swing is observed during the period under study. In the case of the two regions OACM & OCACC, upswing has been observed in two years. For the world as a whole, such upswing has been observed during 1990 and 1993.

As regards "swing-outliers", it varies between very low value of 0.17 to as high as 1. Thus one may conclude that India's CMS faired in favourable direction during the period under study.

The striking feature of the findings is that in most cases upswing is noticed during 1991 - the beginning of the period of economic liberalisation. After 1991, CMS behaved in a positive direction on sustained basis.

**Table II**   Study of upswing and downswing of change in India's import share in the world merchandise trade

| Economic blocs | Upswing behaviour | | Downswing behaviour | |
|---|---|---|---|---|
| | Swing frequency | Swing/ outlier | Swing frequency | Swing/ outlier |
| East African Community | 0 | 0 | 1 | 0.25 |
| Organisation Commune Afri-canine El Malagache | 0 | 0 | 1 | 0.17 |
| Organisation of African Unity | 0 | 0 | 1 | 0.25 |
| Other African Countries | 0 | 0 | 1 | 0.25 |
| North America | 1 | 1 | 1 | 1 |
| South America | 1 | 0.5 | 0 | 0 |
| Caribbean Common Market | 0 | 0 | 2 | 0.4 |
| Central American Common Market | 1 | 0.5 | 2 | 0.67 |
| Other Central American & Caribbean Countries | 2 | 0.67 | 0 | 0 |
| Economic & Social Commission for Asia & Pacific | 1 | 0.17 | 0 | 0 |
| Economic Commission for West Asia | 0 | 0 | 2 | 1 |
| Other Asian & Oceanian, Antarctic and Arctic Region Countries. | 2 | 1 | 0 | 0 |
| European Common Market | 1 | 1 | 2 | 1 |
| European Free Trade Area | 0 | 0 | 1 | 0.33 |
| East European Countries | 0 | 0 | 0 | 0 |
| Other European Countries | 2 | 1 | 0 | 0 |
| All Regions | 0 | 0 | 2 | 1 |

From Table II, we find a somewhat opposite picture when we study the "Change in import share"(CIS) of India in the world market. Downswing has been very common. At least half of all the economic regions in the world has shown downswing. At all regions-level, we observe downswing twice - in 1990 and in 1994. This implies that India's increasing export earning made no impact on the changing world import share. India has to strive very hard for a much higher level of output in its export performance in terms of dollars.

# RECOMMENDATIONS SUMMARISED FROM THE ARTICLES

## (A) Content and Structure of Information

1. The national statistical system should be *adaptive* in its content, i.e. in respect of the data to be generated by the system.

2. The system should be able to produce information at the level of the entire nation, State, district, local self-governing agencies like municipalities, *panchayats*, etc., to help managerial decision-making at all these levels.

3. A short-term economic policy requires data on weekly, monthly and quarterly bases. The data should relate to areas of concern at all levels such as functional groups, commerce and industry, various social and occupational groups (including weaker and disadvantaged sections by sex and age), so that specific programmes may be run and monitored over time for the target groups.

4. In States with rapid urbanisation, urban statistics should be a priority sector. Schemes like publishing Municipal Year-Books should be revived.

5. The NSSO monthly per capita expenditure (MPCE) classification is useful only at the macro-level. NSSO should base the classification on socio-economic characteristics so that results may relate to identifiable target groups.

## (B) Creation of Databases

1. With the huge volume of data already in existence, a menu-driven, user-friendly database can be made available to users throughout the country with the help of a computer network using dedicated P&T line or through satellite. Such a step may also enforce the timely release, and improve the reliability, of data generated by various organisations.

2. A comprehensive central database can be set up at the CSO in respect of data available with the CSO, the NSSO and the Central ministries. The same may be extended to the States in due course. For better dissemination capabilities, this centre should not only be connected with standard networks like the INTERNET or NICNET but should also be equipped with adequate reprographic facilities to make selected data available to officials and data users. The database can be gradually extended to cover administrative data as well.

3. Presentation of geographical database with representative points on maps and charts is a very useful tool for a planner or an administrator. Such geographical databases can be built up with the PIN codes created by the Indian Postal Department. By including PIN code as part of geographical identification of any information, our geographical databases may be easily augmented.

4. At present, there is no comprehensive database on prevalence or incidence of even the major diseases in India. All the data that exist on each disease can be pooled to create such a database. For each major chronic disease in India, one of the academic and government institutions working in the field can serve voluntarily as a nodal centre where all available data and literature on the subject will be collected. The computerised databases of private hospitals and medical insurance companies can supplement these sources. Such data, collected over a period of time, can offer guidance with respect to the changing disease profile of the country and lead to a national health care system.

## (C) Greater Use of Computers

1. Government and semi-government organisations may be modernised with the latest computer and communication systems, while maintaining hardware compatibility among the concerned offices. Further, E-mail and INTERNET facilities should be provided at least to the heads of all concerned statistical offices.

2. The DoS should establish a dedicated satellite-based national computer network so that it need not depend on any outside agency for data processing and data dissemination. There is also a need to modernise data collection systems.

3. So far, computers have been used merely for data processing. They should also be used for design and control of statistical operations.

## (D) Data Collection

1. Orientation should shift from routine data collection and providing fixed-product tables to a demand-led system based on regular interaction between data collection agencies and data users.

2. The possibilities of using data already available should be explored before taking up fresh surveys.

3. The tendency for user agencies in the government to ask for more information, on more items and at a high level of disaggregation needs to be checked as this would adversely affect costs and data quality of surveys. Restraint must also be observed in conducting censuses of various kinds keeping in view the increased resistance from the different field-level functionaries.

4. Refusal of respondents needs to be tackled through (i) reduction in the size of the schedules/ questionnaires by excluding any item which is not considered for generation of tables and (ii) developing alternative techniques to the household sample approach like the quantal method, and exploring new techniques for data collection and processing (e.g. remote sensing in agriculture).

5. In certain cases, the respondent is unable to provide data for the reference period in question but readily reports the information for a shorter period. The investigator then inflates the figures suitably and records the entry for the prescribed reference period. The effect of such adjusted responses could be an area of study.

6. *Data linkages.* If among data elements of a survey or a census, certain key items derived from other surveys are included, it may help in establishing linkages between the fields of data covered by different projects, producing a synergic effect.

7. The NSS consumer expenditure enquiry should use two different questionnaires in two half-samples of households: (i) a detailed Food Schedule followed by an abridged Non-food Schedule and (ii) an abridged Food Schedule followed by a detailed Non-food Schedule. Innovative estimation procedures will have to be evolved to maintain comparability of estimates over time.

8. Possibilities of using palmtop and laptop computers to eliminate the present data entry system for survey data should be explored.

## (E) Sampling Design and Estimation Procedures in the NSSO

1. Changes over time are better measured and better understood if panel data are used by the NSS.

2. Since the NSSO generally undertakes household surveys and surveys on unorganised manufacture and trade, the following method is proposed for allocation of its field posts among the States/UTs. Three different allocations - one in proportion to population and the other two in proportion to number of workers engaged in unorganised manufacture and trade, respectively, may first be

obtained. An ideal allocation may be worked out which might be the simple average of the three allocations or some weighted average of them.

3. For generating efficient estimates for smaller domains, like *tehsils* or districts, (i) strategies involving allocation of samples can be adopted at the design stage, so as to make significant gain in efficiency of district-level estimates at the expense of a small loss of efficiency for a large domain, say, a State, and (ii) a model-based approach involving an efficient use of auxiliary information can be adopted at the estimation stage. These methods are likely to be useful in obtaining reliable district-level estimates with limited resources.

4. Increase of the Central sample size to get reliable estimates for small areas is not practicable. Small area estimates should be obtained by States through decentralisation methods. Sophisticated techniques of small area estimation should be tried on data from Central samples to provide a background for comparison.

5. Non-sampling error is an important component of total error in any large-scale survey. The NSS can benefit by optimising resource allocation with a view to reducing total error, and not just the sampling error.

6. The NSSO should be provided with adequate resources to
   a) have a supporting research wing with the capability to continually keep abreast of the rapidly changing world by utilising some of the survey techniques developed in the post-Mahalanobis era, namely, 'small area estimation', 'randomised response', 'network sampling', 'adaptive sampling', 'remote sensing', etc., and
   b) launch a series of seminars to expose its own *rôles* to scrutiny by external experts.

## (F) Sampling Frames

1. The economic census conducted along with the house-listing operation of population censuses may give counts of enterprises of different types, which may be used to draw samples of follow-up surveys of the economic census. An economic census may be conducted at the mid-point of the period between two population censuses to check the growth/death of enterprises during the intervening period. While carrying out the population censuses or economic censuses, a unique geo-coding structure for all the villages and Enumeration Blocks (EBs) of the country should be devised.

2. At the time of conducting a population census, villages with large populations may be divided into permanent areal segments with identifiable boundaries that are not changed in the subsequent censuses. These segments can be directly selected in the NSS surveys as the first-stage units. This will avoid the present necessity of dividing larger villages into a number of areal segments and surveying only one or two randomly selected segments. The same procedure can be adopted for towns and cities. Also, certain additional auxiliary information for each village or "segment" of a large village can be collected during a population census that can be subsequently used for stratification and sample selection in socio-economic surveys.

## (G) Wider and Better Use of Statistical Methods

1. Analysis of sub-sample results and estimation of errors used to be undertaken earlier. This practice has been discontinued. It must be revived.

2. Well-coordinated projects should be undertaken with various establishments in the government sector, the private sector and universities to incorporate current developments in methodology.

3. *Data synthesis*. Required information is often not available from a single source. Research is needed to develop appropriate techniques of analysis of dispersed data to create the necessary information-base.

4. Statistical offices should tackle technical problems not intuitively, as they do now, using a "common sense" approach, but by attempting to make use of research results obtained elsewhere, or undertaking an independent research investigation to solve the problem.

5. Outside India, techniques are emerging fast for tackling both 'unit' and 'item' non-response by dint of 'imputation' and 'weighting adjustments' and other requisite remedial measures. The NSSO has not so far used diverse methods like balanced half-sampling, balanced repeated replication, jack-knifing and bootstraps, commonly employed devices in estimating standard errors and setting confidence limits to estimated parameters. This needs to be taken up at the earliest.

**(H) Coordination and Standardisation of Statistical Work**

1. It is highly desirable that the information provided by various agencies conform to some specific standards. There should be clearly defined *rôles* for various persons/organisations involved in the task of providing information.

2. For effective coordination and standardisation of statistical activities in the country, the Department of Statistics should be brought under the Cabinet Secretariat. This step would enhance the Central control on the various data sources originating in different departments and ministries. A very senior statistician of the CSO, for example, the officer in the administrative hierarchy just below the Director General, CSO, should be given full responsibility for coordination of 'all statistical activities' of the country.

3. For ensuring uniformity in concepts, definitions and methodologies, the CSO may consider the setting up of subject panels for all important areas. The subject panels, apart from having representatives from the CSO/DoS, should have subject matter specialists from administrative agencies, universities, research institutions and user organisations. The NABS can oversee the working of these subject panels.

4. Dedicated Central/State statistical agencies may be made fully responsible for providing the strategic information. Statistical agencies in ministries should be responsible for management information and individual executing agencies for operational information. A representative of the Central statistical agency present in each functional ministry can play the coordinating *rôle* between the ministry and the Central agency and guide the ministry in developing a suitable information collection system.

5. For subjects under the Concurrent List, sharing of data by Central ministries with States should be as mandatory as it is for the States to furnish data to the Centre.

6. In order to ensure adoption of standard concepts, definitions and methodology and avoidance of duplication of work, it should be made obligatory on the part of statistical agencies to obtain clearance from the NABS before launching any major statistical operation funded by the government.

7. As far as possible, one set of data on a particular subject should be collected/released by a single agency. When more than one agency are involved, they should make all efforts to achieve uniformity in concepts, definitions and survey methodology.

8. Studies should be undertaken to prescribe adjustments to resolve differences in data arising from alternative sources.

9. The government should make special efforts to standardise the procedure of poverty measurement. Since the studies of poverty measurement are mostly based on NSS consumer expenditure data, and the NSS survey on consumer expenditure is now conducted every year, a regular series of all-India and State-specific poverty indices should be brought out in an official publication.

10. For better coordination of statistical activities within a ministry, it is desirable to have a whole-time statistical adviser at a sufficiently high level who can directly report to the seniormost functionary of the ministry. Such statistical advisers should be professional statisticians with sufficient experience in official statistics.

## (I) Reduction in Time Lag

1. Delays in publishing results have plagued the DoS for a long time, though of late there has been considerable improvement in this respect. An advance calendar for release of data, at least in respect of important statistical activities, may help reduce the time lag. The calendar should be prepared and brought out through a public notification at the beginning of each calendar year. A close fortnightly or monthly monitoring system may be introduced to assess the progress of work.

2. Every statistical agency should devise a system to provide a series of quick provisional estimates and summary results. In addition, suitable indicators should be developed and made available in each sector for periodical monitoring and policy adjustment. For important surveys where processing of entire data causes delay in release of results, advance tables based on sub-samples may be generated for the use of planners and policy-makers.

3. The Tabulation Plan of NSS surveys must be simple so that on-demand tables can be produced by both the Centre and the States.

4. The NSS State sample estimates are not available on time for the majority of the States. The data processing capabilities of the DPD, NSSO may be extended to the States which are lagging behind in this respect. This will help in pooling the Central and State sample data for generating improved estimates or releasing estimates at a further disaggregated level.

5. Use of computer packages for generation of tables should be preferred to writing fresh programs.

## (J) Legal Measures

1. For proper coordination/control of statistical activities and for ensuring timely generation of quality statistics, optimum utilisation of resources and avoidance of duplication of work, the DoS requires to be amply strengthened with appropriate statutory authority.

2. There should be an appropriate Statistical Act vesting sufficient administrative powers in the statistical authorities for collection of data, with provision for suitable penalty for refusal to provide information.

3. An appropriate legislation has to be framed whereby the different agencies collecting data regularly or on *ad hoc* bases may be enjoined to adopt a fixed set of accepted definitions and procedures. In addition, the agencies should be statutorily required to provide the collected data to a constituent body of the National Statistical System.

## (K) National Accounts

1. Efforts should be made to study the margins of uncertainty of the estimates of at least the important parameters in the published national accounts statistics.

2. With the liberalisation and globalisation of the economy, there have come into existence a large number of enterprises earning income through new types of activity. Surveys should be undertaken to produce data on these new kinds of income.

3. Existing data on depreciation, estimated on the basis of estimated/assumed life of assets, are not very realistic and there is a need to undertake special censuses and sample surveys of the fixed assets of various industries.

## (L) Industrial and Trade Statistics

1. It is desirable that all the registered industrial units (covered under the ASI) are surveyed once in five years. There should also be monthly surveys whose results should be released within a month of completion of the survey. With the availability of monthly estimates, the compilation and release of the Index of Industrial Production would no longer be necessary in its present form.

2. *Informal sector.* There is a need to conduct methodological studies for collecting dependable information from the unorganised sector units. To estimate the value added in trading activities, instead of collecting both expenditure and sales data, the method of collecting data on purchase and trade-margin directly can be explored. Various sub-sectors of the informal sector, particularly the newly-emerging areas, may be properly identified and surveys conducted for each of the sub-sectors to obtain a proper estimate of the GDP.

3. With some coordination with the power supply authorities, it should be possible to prepare a good frame, useful in collection of industrial production data, of all industrial units using electricity.

## (M) Population Census and Demographic Data

1. Till 1961, the census used to give the actual count of non-agricultural workers by detailed occupational classification at the district level. Since then, this practice has been discontinued. It should be revived.

2. With increasing emphasis on decentralisation of planning, the sample size of national demographic surveys, such as the National Family Health Survey, will have to be increased to provide estimates at the district level. Increase in sample size will increase the workload and may increase the non-sampling error unless the quality of supervision during data collection and processing is improved. One of the ways of ensuring better quality of data is to distribute the survey work-load. While health workers can be engaged in field work of surveys which will estimate demographic parameters, persons working in development and agriculture departments may be asked to conduct household enquiries on employment, occupation and organised sector work-force.

3. There is a need to strengthen the system of sample registration on a regular basis. Efforts are required, and are being made, to improve the coverage in registration of births, deaths and marriages. Provision of electronic equipment at the grassroots level will certainly improve the collection, compilation and transmission of data, and their dissemination at the district or block level.

## (N) Agricultural Census and Livestock Census

1. The agricultural and livestock censuses should be integrated. Separation of the two subjects (land and livestock holdings) is quite unrealistic in our economy, where cultivation of land and animal husbandry are inseparable parts of a single integrated economic activity.

2. The agricultural census, being only a census of land records, can gather much more information without additional cost: for example, the rate of basic land revenue, which is the only available

indicator of inherent productivity of land. Also, more tabulation may usefully be done of the data collected.

## (O) New Areas

1. At present, the requirements of data on emerging areas like natural resource management, ecological changes, industrial metabolism, environmental impacts of urbanisation, dumping of foreign consumer goods, disposal of hazardous wastes, etc., are met by piecemeal studies undertaken by the state or private agencies. The National Statistical System has to take up the task of synthesising the findings of these isolated studies through statistical methods and develop workable databases on the emerging topics not currently covered by official statistics.

2. Measurement of factors associated with *environmental degradation and ecological imbalances* needs statistical survey work, e.g. environmental mapping of intensity of pollution as done in the USA or Canada.

3. Increased industrial and domestic need for energy and mounting consumption of gasoline and diesel products have caused an energy crisis, which has been labelled as a primary concern. An Energy Plan is now a must. The success of such a Plan depends on statistical feedbacks. The collection of the necessary *energy and environment statistics* is a gigantic task which agencies having the necessary statistical expertise, such as the NSSO, should take up immediately.

4. *New diseases* such as AIDS require concentrated attention and research. Public health has emerged as a very important discipline. But the health research institutes which have come up are not the best agencies for monitoring the statistical methodology employed in health research. It would be best if the NSSO could, in collaboration with medical and health care personnel, devise appropriate statistical methods for all statistical aspects of health research from data collection to analysis of results.

5. In the area of human genome research, many of the current human geneticists in India have been collaborating with foreign experts. Often such joint research ventures lead to transporting human DNA samples out of the country whose documentation after the conclusion of the initial study is not known. Since such samples can be preserved over a long period of time with the current technology, the question remains as to who has the right to any future intellectual information on these samples. The Department of Statistics with the help of the Department of Science and Technology should formulate mandatory means of having a *statistical inventory of such research*.

6. Floriculture is an industry in which India can, by virtue of the progress it has made in biotechnology and the variety and uniqueness of its flora, earn 25 to 30 times more foreign exchange than from its other agricultural products. Documented *statistics of the uniqueness of Indian resources* may enable the Indian intellectual rights to its floral heritage to be preserved in case the industry is developed.

7. Perhaps the official statistical system may be able to build means of enforcing accountability of public officials through which *evidence of corruption* at its early stage may be detected and brought to attention. This may require an innovative strategy of monitoring and analysis of expenditure on governmental projects where the individual privacy rights of a citizen should be acknowledged and honoured, yet misuse and misappropriation of authority can be rectified and controlled at an early stage.

## (P) Research in Statistics

1. The CSO should maintain a core staff in each discipline which should also be engaged in research. They should be retained and promoted on their posts. Personnel from universities and

research institutions should be inducted at higher levels for a period of two to three years to give impetus and objectivity.

2. The NSSO should be provided with adequate resources to help its own journal *Sarvekshana* acquire the status of a readable scientific journal with original contributions on methodological issues.

3. The university system does not provide much incentive for applied research in statistics. The DoS in its own interest may step in and sponsor or support applied research in statistics in areas that are important to the DoS, such as the following:

(i) sampling design for multipurpose large-scale surveys; (ii) estimation of sampling and non-sampling errors; (iii) use of auxiliary information from past surveys, or other sources, to improve current estimates; (iv) logical basis for imputation of missing observations; (v) generalised software for survey data processing; (vi) management and control of large-scale surveys to avoid time and cost over-runs; (vii) small area estimation; (viii) use of remote sensing technology in surveys; (ix) identifying the poor; (x) short-term forecasting; (xi) modelling the Indian economy.

## (Q) Training of Statistical Personnel

1. The DoS should arrange frequent training programmes for statistical officers on new methodologies in sample surveys, time-series analysis and forecasting, computer applications and other relevant topics.

2. Training imparted to statistical personnel should be task- and subject-oriented.

3. The Training Division of the DoS should interact with universities and applied statisticians for drawing up meaningful syllabi for the purpose of training for various levels of ISS officers.

## (R) Statistical Education

1. Currently, there is a lack of balance among the three wings of the Indian statistical system, viz. academic activities, applied research and system of official statistics, the first getting the major attention. For restoring the balance, there should be in our training/teaching programmes greater stress on data collection, data analysis and proper interpretation of the findings. 'Practical work', preferably based on live data, should form a major part of the programme.

2. Present statistics curricula at the Bachelor's /Master's level are not adequate for official statisticians. Greater emphasis is needed on application of statistics in planning and administration, statistical modelling and use of computers.

3. Students of statistics at the post-graduate level need to specialise in a chosen subject like economics, public health, agriculture, mathematics and marketing - to be taught by experts in the subject. A course-ending project work in the chosen subject should be part of the curriculum. This kind of a course would produce better-equipped professionals.

## (S) Dissemination of Results

1. The NSSO should revive its earlier practice of releasing State-level data in separate volumes.

2. The DoS should bring out popular publications at regular intervals to apprise the data users about the structure of the Indian Statistical System, current data availability, current controversies relating to quality of statistics and methodology for computing specific indices. The DoS should also lay down a procedure for making raw data available to researchers.

3. *Error estimates*. The DoS has seldom provided to the user any idea about the margin of error of the statistics. It is an ethical necessity to provide to the user an estimate, however crude, of the margin of error of the statistics, measured by the extent of sampling and non-sampling errors.

4. The NSSO should spend more time and resources on press briefings and distribution of summary reports.

**(T) Personnel Policy**

1. The selection test for recruitment to the ISS should put greater emphasis on practical topics than on abstract theoretical aspects.

2. It is desirable that ISS officers who have served for about 15 years in the statistical system be allowed to specialise in the area where they feel themselves most capable of making significant contributions on the basis of their experience.

## List of Subjects Covered and Reports Brought out for the NSS Rounds 36 to 50

| Report No. | Title of the Report |
|---|---|
| (1) | (2) |

**Round No. 36 (July' 81 to Dec.' 81)**
**Survey of Disabled Persons (Sch. 26)**

| | |
|---|---|
| 305 | *Report on Survey of Disabled Persons* |
| 337 | *Characteristics of Disabled Persons* |

**Round No. 37 (Jan. '82 to Dec.' 82)**
**Land Holding & Livestock Holdings (Sch. 18.1)**
**Debt & Investment (Sch. 18.2)**

| | |
|---|---|
| 318 | *Assets and Liabilities of Rural and Urban Households (States & All-India)* |
| 322 | *Some Aspects of Indebtedness of Rural and Urban Households (States & All-India)* |
| 328 | *Some Aspects of Indebtedness of Rural and Urban Households (Pt - II, States & All-India)* |
| 330 | *Some Aspects of Household Ownership Holdings (States & All-India)* |
| 331 | *Some Aspects of Operational Holdings* |
| 334 | *Broad Features of Capital Expenditure Incurred by Rural & Urban Households during '81 - '82 (States & All-India)* |
| 338 | *Estimates of Live-stock and Agricultural Implements Classified by Household Operational Holdings* |
| 340 | *Extent & Nature of Borrowings and Repayments of Rural & Urban Households during July '81-June '82 (States & All-India )* |
| 344 | *Features of Operational Holdings during Kharif & Rabi Seasons of the Agricultural Year 1981-82 (States & All-India, Rural sector)* |
| 349 | *Assets, Liabilities & Capital Expenditure of Urban Households by Population Size Class of Towns (All-India)* |

**Round No. 38 (Jan. '83 to Dec.' 83)**
**Consumer Expenditure (Sch. 1.0)**
**Employment & Unemployment (Sch. 10)**

| | |
|---|---|
| 315 | *Key Results of Last Three Quinquennial NSS Enquiries on Employment & Unemployment* |
| 319 | *Report on the Third Quinquennial Survey on Consumer Expenditure* |
| 332 | *Pattern of Consumer Expenditure of S.C. & S.T.. Households.* |
| 333 | *Pattern of Consumer Expenditure of Households Self-employed in Agriculture and of Agricultural and Rural Labour Households* |

| Report No. | Title of the Report |
|---|---|
| (1) | (2) |
| 336 | *Report on Sources of Drinking Water and Energy Used for Lighting and Cooking* |
| 339 | *Tables with Notes on Particulars of Dwelling Units* |
| 341 | *Report on the Third Quinquennial Survey on Employment and Unemployment : All-India* |
| 341/1 | *Report on the Third Quinquennial Survey on Employment and Unemployment : Andhra Pradesh* |
| 341/2 | *Report on the Third Quinquennial Survey on Employment and Unemployment : Bihar* |
| 341/3 | *Report on the Third Quinquennial Survey on Employment and Unemployment : Kerala* |
| 341/4 | *Report on the Third Quinquennial Survey on Employment and Unemployment : Maharastra* |
| 341/5 | *Report on the Third Quinquennial Survey on Employment and Unemployment : Karnataka* |
| 341/6 | *Report on the Third Quinquennial Survey on Employment and Unemployment : Madhya Pradesh* |
| 341/7 | *Report on the Third Quinquennial Survey on Employment and Unemployment : Orissa* |
| 341/8 | *Report on the Third Quinquennial Survey on Employment and Unemployment : Rajasthan* |
| 341/9 | *Report on the Third Quinquennial Survey on Employment and Unemployment : Tamil Nadu* |
| 341/10 | *Report on the Third Quinquennial Survey on Employment and Unemployment : Uttar Pradesh* |
| 341/11 | *Report on the Third Quinquennial Survey on Employment and Unemployment : West Bengal* |
| 341/12 | *Report on the Third Quinquennial Survey on Employment and Unemployment : Gujarat* |
| 341/13 | *Report on the Third Quinquennial Survey on Employment and Unemployment : Haryana* |
| 341/14 | *Report on the Third Quinquennial Survey on Employment and Unemployment : Punjab* |
| 341/15 | *Report on the Third Quinquennial Survey on Employment and Unemployment : Assam* |
| 341/16 | *Report on the Third Quinquennial Survey on Employment and Unemployment: Himachal Pradesh* |
| 341/17 | *Report on the Third Quinquennial Survey on Employment and Unemployment: Jammu & Kashmir* |
| 341A | *Employment and Unemployment Situation of S.C. & S.T. Population during Early Eighties* |
| 341B | *Employment and Unemployment Situation in Cities & Towns during Early Eighties* |
| 348 | *Per Capita and Per Consumer Unit per Diem Intake of Calorie, Protein and Fat and Perception of the People on Adequacy of Food* |

| Report No. | Title of the Report |
|:---:|:---:|
| (1) | (2) |
| 352 | *Monthly per Capita Consumption of Cereals for Various Sections of Population* |
| 353 | *Level of Nutritional Intake of Population Distributed over Different Expenditure Classes* |
| 356 | *Tables with Notes on the Effect of Adjustment of Energy Intake for Meals Consumed Free and Meals served to Others* |
| 357 | *Tables with Notes on Distribution of Households by Monthly Per Capita Expenditure Class after Adjustment (Considering Annual Expenditure)* |
| 387 | *Distribution of Households and Persons by Household Monthly Per Capita Expenditure for Different Calorie Intake Levels* |

**Round No. 39 (Jan.' 84 to Jun. '84)**
**Population, Births & Deaths Enumeration (Sch. 12)**
**Population, Births & Deaths Enumeration & Re-enumeration (Sch. 12.1)**
**Village Particulars (Sch. 3.1)**

| | |
|:---:|:---|
| 350 | *Effect of Drought on the Pattern of Employment and Unemployment: a Comparative Study of the Survey Results of NSS 38$^{th}$ and 39$^{th}$ Rounds* |
| 351 | *Population, Birth & Death Rates* |
| 354 | *Fertility Rates in India* |

**Round No. 40 (July '84 to Jun. '85)**
**Unorganised Manufacture: Non-directory Establishments and**
**Own Account Enterprises (Sch. 2.2B)**

| | |
|:---:|:---|
| 363/1 | *Tables with Notes on Survey of Unorganised Manufacture : Non-directory Establishments and Own Account Enterprises - All-India* |
| 363/2 | *Tables with Notes on Survey of Unorganised Manufacture : Non-directory Establishments and Own Account Enterprises - States (Vol. I & II)* |

**Round No. 41 (July '85 to Jun. '86)**
**Trade: Non-directory Establishments & Own Account Enterprises (Sch. 2.41B)**

| | |
|:---:|:---|
| 366/1 | *Tables with Notes on Own Account Trading Enterprises and Non-directory Establishments: All-India* |
| 366/2 | *Tables with Notes on Own Account Trading Enterprises and Non-directory Establishments: States (Vol. I & II)* |

| Report No. | Title of the Report |
|---|---|
| (1) | (2) |

### Round No. 42 (July '86 to Jun. '87)
**Maternity, Child Care, Family Planning & Utilisation of Public Distribution System (Sch. 25.1)**
**Participation in Education (Sch. 25.2)**
**Survey on Persons Aged 60 Years & above (Sch. 27)**
**Survey on Ex-armed Force Personnel (Sch. 28)**
**Consumer Expenditure (Sch. 1.0)**

| | |
|---|---|
| 355/1 | *Tables with Notes on Consumer Expenditure for Assam, Bihar, Manipur, Meghalaya, Mizoram, Nagaland, Orissa, Rajasthan, Sikkim, Tamil Nadu, Tripura, U.P., West Bengal and Pondicherry* |
| 355/2 | *Tables with Notes on Consumer Expenditure for All-India, Andhra Pradesh, Gujarat, Haryana, H.P., J&K, Karnataka, Kerala, M.P., Maharashtra, Punjab, A & N Islands, Chandigarh, D & N Haveli, Delhi, Goa, Daman & Diu, Lakshadweep* |
| 361 | *A profile of Households & Population by Economic Class and Social Group and Availability of Drinking Water, Electricity and Disinfection of Dwellings* |
| 362 | *Utilisation of Public Distribution System* |
| 364 | *Morbidity and Utilisation of Medical Services* |
| 365/1 | *Participation in Education : All-India* |
| 365/2 | *Participation in Education : Major States (Vol. I)* |
| 365/2 | *Participation in Education : Major States (Vol. II)* |
| 367 | *Socio Economic Profile of the Aged Persons* |
| 368 | *Child & Maternity Care* |
| 369 | *Utilisation of Family Planning Services* |

### Round No. 43 (July '87 to Jun. '88)
**Consumer Expenditure (Sch. 1.0)**
**Employment & Unemployment (Sch. 10)**
**Travel Habits of Households (Sch. 21)**

| | |
|---|---|
| 371A | *Employment & Unemployment Situation of Scheduled Tribe and Scheduled Caste Population during Late Eighties* |
| 371B | *Employment and Unemployment Situation in Cities & Towns during Late Eighties* |
| 372 | *Pattern of Consumer Expenditure* |
| 373 | *Particulars of Dwelling Units (SS-I)* |
| 374 | *Pattern of Consumption of Cereals, Pulses , Tobacco & Some Other Selected Items* |

| Report No. | Title of the Report |
|---|---|
| (1) | (2) |

**Round No. 44 (July '88 to Jun. '89)**
**Village Characteristics (Sch. 3.1)**
**Level of Living of Tribals (Sch. 29.1)**
**Economic Activities of Tribals (Sch. 29.2)**
**Particulars of Migration & Ownership of Land by Non-tribals in Tribal Areas (Sch. 29.3)**
**Housing Condition (Sch. 1.2)**
**Particulars of Construction (Sch. 1.4)**

| 370/1 | *Tables with Notes on Consumer Expenditure (All-India), Semi Rounds 1* |
| 370/2 | *Tables with Notes on Consumer Expenditure (All-India), Semi Rounds 1 & 2* |
| 375 | *A Report on Some Important Characteristics of Villages in Tribal Areas* |
| 376 | *Report on Housing Condition* |
| 377 | *Report on Building Construction* |
| 378 | *A Report on Living Conditions of Some Major Tribes of India* |
| 379 | *An Economic Profile of Some Major Tribes of India* |
| 380 | *A Report on Living Conditions of Tribals and Non-tribals of Tribal Areas* |
| 389 | *A Report on an Exploratory Survey of Living Conditions of Tribals of Nagaland* |

**Round No. 45 (July '89 to Jun. '90)**
**Unorganised Manufacture: Non-directory**
**Establishments & Own Account Enterprises (Sch. 2.2)**
**Consumer Expenditure (Sch. 1.0)**

| 381 | *Tables with Notes on the Third Annual Survey on Consumer Expenditure* |
| 396/1 | *Tables with Notes on Survey of Un-organised Manufacture : Non-directory Establishments & Own Account Enterprises, Part-I: All-India* |
| 396/2 | *Tables with Notes on Survey of Un-organised Manufacture : Non-directory Establishments & Own Account Enterprises, Part-II: States* |

**Round No. 46 (July '90 to Jun. '91)**
**Trade: Non-directory Establishments and Own Account Enterprises (Sch. 2.41.2)**
**Consumer Expenditure (Sch. 1.0)**

| 386 | *Tables with Notes on Fourth Annual Survey on Consumer Expenditure and Employment - Unemployment* |
| 403 | *Small Trading Units in India* |
| 403/1 Vol.-I | *State Level Results on Small Trading Units in India* |
| 403/1 Vol.-II | *State Level Results on Small Trading Units in India* |

| Report No. | Title of the Report |
|---|---|
| (1) | (2) |

**Round No. 47 (July '91 to Dec. '91)**
**Village Facilities (Sch. 3.1)**
**Survey of Disabled Persons (Sch. 26)**
**Developmental Milestones of Children (Sch. 26.1)**
**Literacy & Culture (Sch. 30)**
**Consumer Expenditure (Sch. 1.0)**

| | |
|---|---|
| 388 | *Tables with Notes on Fifth Annual Survey on Consumer Expenditure & Employment - Unemployment* |
| 392 | *Availability of Some Education and Culture Related Facilities in Indian Villages* |
| 393 | *A Report on Disabled Persons* |
| 394 | *Literacy in India* |
| 395 | *Participation in Cultural Activities* |

**Round No. 48 (Jan. '92 to Dec. '92)**
**Land & Livestock Holdings (Sch. 18.1)**
**Debt & Investment (Sch. 18.2)**
**Consumer Expenditure (Sch. 1.0)**

| | |
|---|---|
| 397 | *Household Consumer Expenditure and Employment Situation in India* |
| 399 | *Some Aspects of Household Ownership Holdings* |
| 407 | *Operational Land Holdings in India, 1991-92, Salient Features* |
| 408 | *Livestock and Agricultural Implements in Household Operational Holdings, 1991-92* |
| 414 | *Seasonal Variation in the Operation of Land Holdings in India, 1991-92* |
| 419 | *Household Assets and Liabilities as on 30.6.91* |

**Round No. 49 (Jan. '93 to Jun. '93)**
**Particulars of Slums (Sch. 0.21)**
**Housing Condition & Migration (Sch. 1.2)**
**Consumer Expenditure (Sch. 1.0)**

| | |
|---|---|
| 400 | *Household Consumer Expenditure and Employment Situation in India* |
| 417 | *Slums in India* |
| 429 | *Housing Conditions in India* |

| Report No. | Title of the Report |
|------------|---------------------|
| (1) | (2) |

**Round No. 50 (July '93 to Jun. '94)**
**Household Consumer Expenditure (Sch. 1.0)**
**Employment and Unemployment (Sch. 10)**

| | |
|------------|---------------------|
| 401 | *Key Results on Household Consumer Expenditure, 1993-94* |
| 402 | *Level and Pattern of Consumer Expenditure* |
| 404 | *Consumption of Some Important Commodities in India* |
| 405 | *Nutritional Intake in India* |
| 406 | *Key Results on Employment and Unemployment* |
| 409 | *Employment and Unemployment in India, 1993-94* |
| 410/1 | *Dwellings in India* |
| 410/2 | *Energy Used by Indian Households* |
| 411 | *Employment and Unemployment Situation in Cities and Towns, 1993-94* |
| 412 | *Economic Activities and School Attendance by Children of India, 1993-94* |
| 413 | *Sources of Household Income in India, 1993-94* |
| 415 | *Reported Adequacy of Food Intake in India, 1993-94* |
| 416 | *Participation of Indian Women in Household Work and Other Specified Activities, 1993-94* |
| 418 | *Unemployed in India, 1993-94, Salient Features* |
| 422 | *Differences in Level of Consumption among Socio-Economic Groups* |
| 423 | *IRDP Assistance and Participation in Public Works, 1993-94* |
| 424 | *Ownership of Livestock, Cultivation of Selected Crops and Consumption Levels* |
| 425 | *Employment and Unemployment Situation Among Social Groups in India, 1993-94* |
| 426 | *Use of Durable Goods by Indian Households, 1993-94* |
| 427 | *Consumption of Tobacco in India, 1993-94* |
| 428 | *Wages in Kind, Exchange of Gifts and Expenditure on Ceremonies and Insurance in India, 1993-94* |

## Publications of the Department of Statistics

| Sl. No. | Name of publication | Periodicity |
|---|---|---|
| 1. | *Monthly Abstract of Statistics* | Monthly |
| 2. | *Statistical Abstract, India* | Annual |
| 3. | *Statistical Pocket Book, India* | - do - |
| 4. | *National Accounts Statistics* | - do - |
| 5. | *Annual Survey of Industries - Summary Results: Factory Sector* | - do - |
| 6. | *Supplement to Annual Survey of Industries* | - do - |
| 7. | *Survey Results - Factory Sector by State & Industry* | - do - |
| 8. | *Annual Survey of Industries, Detailed Results (in 16 volumes)* | - do - |
| 9. | *Selected Socio-Economic Statistics* | - do - |
| 10. | *Women and Men in India* | - do - |
| 11. | *Energy Statistics* | - do - |
| 12. | *Compendium of Environment Statistics* | - do - |
| 13. | *Sarvekshana* | Quarterly |
| 14. | *Statistical System in India* | *ad hoc* |
| 15. | *Sample Survey of Current Interest in India* | Annual |
| 16. | *Guide to Official Statistics* | *ad hoc* |
| 17. | *National Accounts Statistics: Survey and Methodology* | *ad hoc* |
| 18. | *Input-output Transactions* | *ad hoc* |

## Some Important Articles Using DoS and RGI Data Published in Economic and Political Weekly during July 1993 to March 1998

| Author | Title | Vol. | No. | Pages |
|---|---|---|---|---|
| Acharya, Paramesh | Problems of Universal Elementary Education | 29 | 49 | 3098 - 3105 |
| Agnihotri, S. B. | Missing Females: A Disaggregated Analysis | 30 | 33 | 2074 - 2084 |
| Arun Kumar, A. V., Vani, B. P. and Vyasulu, Vinod | Structure of Employment as seen from 1981 and 1991 Census - A Preliminary Look | 30 | 38 | 2375 - 2388 |
| Athukorala, P. and Sen, K. | Economic Reforms and Rate of Saving in India | 30 | 35 | 2184 - 2190 |
| Balakrishnan, P. and Pushpagadan, K. | Total Factor - Productivity Growth in Manufacturing Industry: A Fresh Look | 29 | 31 | 2028 - 2035 |
| Basant , Rakesh | Economic Diversification in Rural Areas - Review of Processes with Special Reference to Gujarat | 29 | 39 | A107 - A116 |
| Bhalla, G. S. and Hazell Peter | Foodgrains Demand in India to 2020 | 32 | 52 | A150 - A154 |
| Bhattacharya, B. B. and Mitra Arup | Employment and Structural Adjustment - A Look at 1991 Census Data | 28 | 38 | 1989 - 1995 |
| Chatterjee, Somnath and Mohan, Rakesh | India's Garment Exports | 28 | 35 | M95 - M119 |
| Chaudhuri, Sudip | Public Enterprises and Private Purposes | 29 | 22 | 1338 - 1347 |
| Dandekar, V. M. | Assets and Liabilities of Government of India | 29 | 3 | 111 - 117 |
| Dandekar, V. M. | Role of Economic Planning in India in the 1990s and Beyond | 29 | 24 | 1457 - 1464 |
| Dev, S. M. | Food Security: PDS vs EGS | 31 | 27 | 1752 - 1758 |
| Dhawan, B. D. and Yadav, S. S. | Private Fixed Capital Expenditure in Agriculture - Some Aspects of Indian Farmers' Investment Behavior | 30 | 39 | A103 - A109 |
| Dholakia, B. H. | Functional Distribution of National Income in India | 30 | 4 | 229 - 236 |
| Dholakia, Ravindra H. | Spatial Dimension of Acceleration of Economic Growth in India | 29 | 35 | 2303 - 2309 |
| Dreze, J. and Loh, J. | Literacy in India and China | 30 | 45 | 2868 - 2878 |
| Duggal, R. Nandraj, S. and Vadair, A. | Health Expenditure across States (in two parts). Part I | 30 | 15 | 834 - 844 |
| Duggal, R. Nandraj, S. and Vadair, A. | Health Expenditure across States (in two parts). Part II | 30 | 16 | 901 - 908 |
| Eapen, Mridul | Rural Non - Agricultural Employment in Kerala - Some Emerging Tendencies | 29 | 21 | 1285 - 1296 |

| Author | Title | Vol. | No. | Pages |
|---|---|---|---|---|
| EPW Research Foundation | Economic Reform and Rate of Saving | 30 | 18 & 19 | 1021 - 1041 |
| George, Alex and Nandraj, Sunil | State of Heath Care in Maharashtra - A Comparative Analysis | 28 | 32 & 33 | 1671 - 1683 |
| Ghanekar, Jayanti | Sorry State of Agricultural Wage data : Sources and Methods of Collection | 32 | 19 | 1029 - 1036 |
| Gopalan, C. | Towards Food and Nutrition Security | 30 | 52 | A134 - A141 |
| Gulati, Ashok and Bhide, Shashanka | What do the Reformers have for Agriculture ? | 30 | 18 & 19 | 1089 - 1093 |
| Gulati, Ashok and Sharma, Anil | Agriculture under GATT: What it Holds for India | 29 | 29 | 1857 - 1863 |
| Gumber, Anil | Burden of Injury in India | 32 | 25 | 1478 - 1491 |
| Gupta, S. P. | Economic Reform and its Impact on Poor | 30 | 22 | 1295 - 1313 |
| Hashim, S. R. | Economic Growth and Income Distribution | 33 | 12 | 661 - 666 |
| Indrakanth, S. | Coverage and Leakages in PDS in Andhra Pradesh | 32 | 19 | 999 - 1001 |
| Kannan, K. P. | Levelling Up or Levelling Down? - Labour Institutions and Economic Development in India | 29 | 30 | 1938 - 1945 |
| Kantor, Paola | Informal Sector: Lifting the Shroud | 32 | 40 | 2512 - 2515 |
| Kulkarni, Sumati | Dependence on Agricultural Employment | 29 | 51 & 52 | 3260 - 3262 |
| Kulshreshtha, A. C., Kolli Ramesh and Singh, Gulab | Agriculture in the National Accounts: Gaps in the Information System | 32 | 27 | 1650 - 1655 |
| Kulshreshtha, A. C. and Singh, G. | Domestic Product by Gender in the Framework of 1993 SNA | 31 | 51 | 3330 - 3334 |
| Kundu, A. and Gupta, S. | Migration, Urbanisation and Regional Inequality | 31 | 52 | 3391 - 3398 |
| Kundu, Amitabh | Trends and Structure of Employment in the 1990's | 32 | 24 | 1399 - 1405 |
| Lalvani, Mala | Some Anomalies in Indian Data Sets : Tackling Inconsistencies | 32 | 23 | 1345 - 1349 |
| Maiti, Pradip and Chattopadhyay, Manabendu | Trends in Level of Living in Urban India | 28 | 46 & 47 | 2547 - 2550 |
| Mathew, E. T. | Educated Unemployment in Kerala - Some Socio Economic Aspects | 30 | 6 | 325 - 335 |
| Meenakshi, J. V. | How Important are Changes in Taste? A State-Level Analysis of Food Demand | 31 | 50 | 3265 - 3269 |
| Mishra, S. N and Chand, Ramesh | Public and Private Capital Formation in Indian Agriculture - Comments on Complementary Hypothesis and Others | 30 | 25 | A64 - A79 |
| Mitra, Ashok | Census 1961: New Pathways | 29 | 51 & 52 | 3207 - 3221 |
| Mohanti, K. K. and Padhi, Sakti | Employment Situation of Tribal Population in Orissa - 1981 Census Data | 30 | 29 | 1879 - 1882 |

| Author | Title | Vol. | No. | Pages |
|---|---|---|---|---|
| Mungekar, B. L. | Inter-sectoral Terms of Trade - Issues of Concept and Method | 28 | 39 | A111 - A119 |
| Nagaraj, R. | What has Happened since 1991? Assessment of India's Economic Reforms | 32 | 44 & 45 | 2869 - 2879 |
| Nagaraj, R. | Employment and Wages in Manufacturing Industries - Trends, Hypothesis and Evidence | 29 | 4 | 177 - 186 |
| Nambiar, R. G. and Tadas, G. | Is Trade Deindustrialising India ? | 29 | 42 | 2741 - 2746 |
| Nambissan, G. B. | Equity in Education ? - Schooling of Dalit Children in India | 31 | 16 & 17 | 1011 - 1024 |
| Nayar, K. R. | Housing Amenities and Health Improvement: Some Findings | 32 | 22 | 1275 - 1279 |
| Parthasarathy, G. and Anand, J. | Employment and Unemployment in Andhra Pradesh | 30 | 15 | 811 - 821 |
| Parthasarathy, G. and Nirmala, K. A. | Lakdawala Estimate of Poverty and Targetted PDS | 32 | 16 | 815 - 816 |
| Paul, S. | Unemployment in India - Temporal and Regional Variations | 28 | 44 | 2407 - 2414 |
| Prathasarathy, G. | Unorganised Sector and Structural Adjustment | 31 | 28 | 1859 - 1869 |
| Rajan, S. Irudaya | Heading towards a Billion | 29 | 51 & 52 | 3201 - 3205 |
| Rajeshwari | Gender Bias in Utilisation of Health Care Facilities in Rural Haryana | 30 | 8 | 489 - 494 |
| Ramachandran, V. K. Rawal, V. and Swaminathan, M. | Investment Gaps in Primary Education - A Statewise Study | 32 | 1 & 2 | 39 - 45 |
| Ramaswamy, K.V. | Small-Scale Manufacturing Industries - Some Aspects of Size, Growth and Structure | 29 | 9 | M13 - M23 |
| Rao, J. M. | Manufacturing Productivity Growth - Method and Measurement | 31 | 44 | 2927 - 2936 |
| Ravallion, Martin | Reform, Food Prices and Poverty in India | 33 | 1 & 2 | 29 - 36 |
| Raychaudhuri, B. | Measurment of Capital Stock in Indian Industries | 31 | 21 | M2 - M6 |
| Reddy, V. Ratna and Rathore, M. S. | Bias in Social Consumption - Case of Residential Water in Rajasthan | 28 | 32 & 33 | 1645 - 1648 |
| Saggar, M. and Pan, I. | SCs and STs in Eastern India - Inequality and Poverty Estimates | 29 | 10 | 567 - 574 |
| Saleth, R. Maria and Thangaraj, M. | Distribution Pattern of Lift Irrigation in India - An Analysis by Hydro-Geological Regions | 28 | 39 | A102 - A110 |
| Sanyal, S. K. | Household Financing of Health Care | 31 | 20 | 1216 - 1222 |
| Saradamani, K. | Women, Kerala and Some Development Issues | 29 | 9 | 501 - 509 |
| Sekhar, P. S. | Levels of Morbidity in Andhra Pradesh | 32 | 13 | 663 - 672 |
| Shah, A. M. | Is the Joint Household Disintegrating ? | 30 | 9 | 537 - 542 |

| Author | Title | Vol. | No. | Pages |
|---|---|---|---|---|
| Shanta, N | Estimates for Private Corporate Sector - A Note | 29 | 29 | 1874 - 1876 |
| Shariff, A. | Socio-Economic and Demographic Differentials between Hindus and Muslims in India | 30 | 46 | 2947 - 2953 |
| Sharma, H. R. | Distribution of Landholdings in Rural India, 1953-54 to 1981-82 - Implications for Land Reforms | 29 | 13 | A12 - A25 |
| Sharma, H. R. | Distribution of Landholdings in Rural India 1953 - 54 to 1981 - 82 | 29 | 39 | A117 - A128 |
| Shiva Kumar, A. K. | UNDP's Gender-Related Development Index - A Computation for Indian States | 31 | 14 | 887 - 895 |
| Srinivasan, K. | Sex Ratios: What They Hide and What They Reveal | 29 | 51 & 52 | 3233 - 3234 |
| Suryanarayana, M. H. | Economic Reforms, Nature and Poverty | 30 | 10 | 617 - 624 |
| Suryanarayana, M.H. | Urban Bias in PDS | 29 | 9 | 510 - 512 |
| Swaminathan, Padmini | Where are the Entrepreneurs? | 29 | 22 | M64 - M74 |
| Tendulkar, S. D. and Jain, L. R. | Economic Reforms and Poverty | 30 | 23 | 1373 - 1377 |
| Tilak, J. B. G. | How Free is "Free" Primary Education in India? Part I | 30 | 5 | 275 - 282 |
| Tilak, J. B. G. | How Free is "Free" Primary Education in India? Part II | 30 | 6 | 355 - 366 |
| Unni, Jeemol | Non-Agricultural Employment and Poverty in Rural India | 33 | 13 | A36 - A44 |
| Upendranath, C. | Structural Adjustment and Education - Issues Related to Equity | 28 | 44 | 2415 - 2419 |
| Upendranath, C. Vijayabaskar, M. and Vyasulu, V. | Industrial Growth and Structure - Analysis of Manufacturing Sector in Karnataka | 29 | 48 | M157 - M164 |
| Vaidyanathan, A. | Employment Situation: Some Emerging Perspective | 29 | 50 | 3147 - 3156 |

## Some Important Articles Using DoS and RGI Data Published in Indian Labour Journal During July 1993 to March 1998

| Author | Title | Year | Month |
|---|---|---|---|
| Lakshmidevi, K. R. | An Analysis of Female Labour Force Participation in Kerala | 1995 | October |
| Ram, Kewal | The Profile, Problems and Protection of Inter - State Circular Migrant Labour in India | 1997 | November |
| Roy, P. K. | Skill Development - Its Perspective in Indian Context | 1995 | January |
| Roy, P. K. | A Retrospection of Employment in Original Sector | 1997 | May |

The important publications/reports brought out by the Central Statistical Organisation (CSO) and the National Sample Survey Organisation (NSSO) are available for sale both on magnetic media and in the form of hard copies at reasonable prices. A major portion of the primary data collected/compiled by the CSO and NSSO is also available on magnetic media for sale to individuals/institutions at a nominal price. Publications/ data of the CSO may be obtained from:

The Director General, CSO
Department of Statistics
Sardar Patel Bhavan
Sansad Marg
New Delhi - 110 001

and those of the NSSO from:

The Chief Executive Officer, NSSO
Department of Statistics
Sardar Patel Bhavan
Sansad Marg
New Delhi - 110 001.